Genomic and Immunological Features of Micropapillary Pattern in Lung Adenocarcinoma

Yansong Huo
Tianjin Medical University Cancer Institute and Hospital
Leina Sun
Tianjin Medical University Cancer Institute and Hospital
Jie Yuan
Geneplus
Hua Zhang
Tianjin Medical University Cancer Institute and Hospital
Zhenfa Zhang
Tianjin Medical University Cancer Institute and Hospital
Lianmin Zhang
Tianjin Medical University Cancer Institute and Hospital
Wuhao Huang Huang
Tianjin Medical University Cancer Institute and Hospital
Xiaoyan Sun
Tianjin Medical University Cancer Institute and Hospital
Zhe Tang
Tianjin Medical University Cancer Institute and Hospital
Yingnan Feng
Tianjin Medical University Cancer Institute and Hospital
Huilan Mo
Tianjin Medical University Cancer Institute and Hospital
Zuoquan Yang
Geneplus
Chao Zhang
Geneplus
Zicheng Yu
Geneplus
Dongsheng Yue
Tianjin Medical University Cancer Institute and Hospital
Bin Zhang
Tianjin Medical University Cancer Institute and Hospital
Changli Wang (✉ wangchangli@tjmuch.com)
Tianjin Medical University Cancer Institute and Hospital

Research Article

Keywords: Lung adenocarcinoma, micropapillary subtype, lepidic subtype, whole-exome sequencing

Posted Date: January 10th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1215973/v1

Abstract

Background: Lung adenocarcinoma (LUAD) usually contain heterogeneous histological subtypes, among which the micropapillary (MIP) subtype was associated with poor prognosis while the lepidic (LEP) subtype possessed the most favorable outcome. A more comprehensive analysis involving discovery and public validation cohorts on the two subtypes could better decipher the key biological and evolutionary mechanisms.

Methods: We firstly retrospectively studied the survival status of 286 LUAD patients with different subtypes. MIP and LEP components were micro-dissected for whole-exome sequencing (WES). Shared and private alterations as well as genomic alternation characteristics between the two components were investigated. Four public cohorts containing LEP and MIP samples were further selected for genomic profile comparison, novel therapeutic target investigation and immune infiltration quantification.

Results: LEP and MIP subtypes exhibited largest disease free survival (DFS) in our patients. A total of 2035 SNV and 2757 InDels were identified in the sequenced LEP and MIP components. *EGFR* was found with highest mutation frequency. Distinct biological processes or pathways were involved in the evolutionary of the two components. Besides, analyses on copy number variation (CNV) and intratumor heterogeneity further discovered the possible immunosurveillance escape, the discrepancy between mutation and CNV level ITH and the pervasive DNA Damage Response as well as WNT pathway gene alternations in MIP component. Phylogenetic analysis on 5 pairs of LEP and MIP components further confirmed the presence of ancestral *EGFR* mutations. Through comprehensive analysis in our samples and public cohorts, *PTP4A3*, *NAPRT* and *RECQL4* were identified as novel therapeutic and diagnostic targets in MIP subtype. Immunosuppression prevalence in MIP component was finally confirmed by multi-omics data.

Conclusion: We identified genetic differences responsible for variated prognosis. The subtype evolution trajectory was additionally unraveled. Novel gene targets and the immunological analyses also provided therapeutic suggestions for MIP subtype.

Introduction

Lung cancer is the leading cause of cancer-related death worldwide, and adenocarcinoma is the most common histological type of non-small-cell lung cancer (NSCLC)(1). Most cases of adenocarcinoma are composed of heterogeneous histological subtypes rather than a single one. In the year of 2015, the world health organization (WHO) proposed a novel definition of five lung adenocarcinoma (LUAD) subtypes to address the histologic heterogeneity, including the lepidic (LEP), acinar (ACI), papillary (PAP), micropapillary (MIP), and solid (SOL) pattern types. More detailedly, as a poorly differentiated, high-grade tumor, the MIP histological subtype has been repeatedly reported to be a negative prognostic factor. Patients presenting MIP are prone to have lymphovascular invasion and pleural invasion, as well as lymph node or intrapulmonary metastasis after surgical resection(2). Meanwhile, previous studies indicated that patients with LEP growth pattern exhibited less aggressive behavior and had most favorable outcome among the predefined subtypes(3). Both the presence of LEP and absence of MIP growth patterns served as predictors of favorable disease-free survival (DFS).

Aiming for the elucidation of the mechanisms beyond tumorigenesis and malignance discrepancy, several studies were conducted to evaluate the molecular and genetic features of LUAD subtypes, especially on MIP and LEP. As for MIP subtype, though positive staining of E-cadherin and beta-catenin were found, a recent study observed the disruption of catenin-cadherin complex in MIP(4), which possibly contributed to its poor intercellular adherence. Besides, the cytoplasmic accumulation of beta-catenin induced the canonical WNT-beta-catenin signaling pathway in MIP. The cell proliferation and migration induced by canonical WNT pathway may account for the histological characteristics of MIP patients. As for the genetic level, MIP/SOL tumors had significantly higher tumor mutation burden (TMB) and fraction of genome altered than other LUAD subtypes. Key oncogenes *BRAF* and *EGFR* were found with higher mutation frequency in LUAD with MIP in multiregional and multiracial cohorts(5). The gene and protein level of *c-MET* were also found elevated in MIP and patients

with poor prognosis(5, 6). Although dysregulated oncogenes associated with poorer prognoses of MIP-predominant LUAD were identified, there remains key mechanisms uncharacterized. For example, the genetic association between subtypes and evolutional trajectory of the relatively malignant MIP subtype were scarcely discussed.

Additionally, therapeutic options including surgical resection, chemotherapy, and targeted therapy have been proposed recently for LUAD with MIP dominance. However, these regimens still have non-negligible limitations. Emerging evidences denoted the limited resection may not be the optimal surgical approach for MIP patients due to observed postoperative recurrence, whereas adjuvant chemotherapy can mostly contribute to early-stage MIP-positive patients(6). Besides, though sporadic mutational targets were identified in previous studies, the development of the effective therapy is burdensome. Not to mention the inescapable acquired resistance to tyrosine kinase inhibitor (TKI) like *EGFR* TKIs. Noticing the recent emergence of lung cancer immunotherapy, studies assessing the efficacy of immune-related therapies on MIP-predominant LUAD had emerged. Considering the abundance of programmed death-ligand 1 (PD-L1) and programmed cell death protein 1 (PD-1) as well as the tumor immunological microenvironment crucially influence the immunotherapy effectiveness, Francois et al. found the significant differences in PD-L1 expression level between LUAD histological patterns(7) while Zhang et al. detected higher CD4+ and CD8+ T cell infiltrations as well as increased PD-L1 abundance through the immunohistochemistry staining(8). Regarding the fact that both the studies focused on restricted components of the tumor microenvironment (TME), A more comprehensive analysis of the variation of TME in specific LUAD subtypes could fill the gap of optimal treatment determination, especially for MIP patients.

To address the limitations mentioned above, we retrospectively reviewed 286 patients with different histological subtype-predominance and compared their survival difference. Patients simultaneously possess MIP and LEP components were further selected for whole-exome sequencing (WES) on both LEP and MIP components and the genetic differences responsible for variated prognosis as well as the subtype-level genetic association were investigated. Multi-cohort analyses further discovered the genes specifically altered in MIP or LEP as novel therapeutic targets as well as the extent of immune infiltration. Our results expanded the evolutional cognition between the LUAD subtypes and offered therapeutic suggestions for MIP patients.

Materials And Methods

Patient selection and histopathologic subtyping

We retrospectively reviewed patients diagnosed with LUAD at Tianjin Cancer Hospital from 2011 to 2014. Patients underwent tumor resection, in which the MIP component exceeded 5% of area size were primarily selected. Later those receiving pre-surgery anticancer treatment, with stage IV disease or other malignancy were further excluded. A total of 286 patients passed the selection criteria and the resected tumors were restaged according to the eighth edition of American joint committee on Cancer TNM staging system for lung cancer. As for the LUAD histological subtyping, the formalin-fixed paraffin-embedded (FFPE) samples were firstly stained with hematoxylin and eosin (H&E) and reviewed by 2 pathologists. The percentage of each histological component was further calculated in 5% increment and the most dominant pattern was recorded.

Clinical data on all patients were extracted from medical records and follow-up information (as of June 2019) was obtained, including disease-free survival (DFS) and overall survival (OS). This study was approved by our institutional review board. Written informed consent was obtained from all patients.

Sample laser-capture micro-dissection and high-throughput sequencing data generation

8 FFPE samples from LUAD patients who underwent surgery at Tianjin Cancer Hospital between 2018 and 2020 were micro-dissected to isolate MIP and LEP components using NIKON ECLIPSE TI2, Japan microscope. More detailedly, 20 FFPE slides were cut in 10 um thick sections, baked for 1h at 60℃, stained with H&E and immersed in xylene. MIP and LEP areas

were circumscribed electronically under the microscope and collected into centrifugation tube with adhesive-cap after ablating by cold laser.

Later genomic DNA of the 5 pairs of MIP and LEP components plus one MIP component passing initial quality control was extracted by Maxwell 16 FFPE Plus LEV DNA purification kit and fragmented by an ultra-sonicator UCD-200 (Diagenode, Seraing, Belgium) with length-based selection through Hieff NGS DNA selection beads. DNA quantity was assessed by Qubit 2.0 Fluorometer with Quanti-IT dsDNA HS Assay Kit (Thermo Fisher Scientific, MA, USA). The sequencing libraries were further constructed by a custom 53M whole-exon capturing probe (IDT, IA, USA). Geneplus-2000 sequencing platform (Geneplus, Beijing, China) further sequenced the libraries in a 100bp paired-end manner.

Mutation calling, somatic copy number alteration detection and mutational signature analysis

Raw sequencing data was primarily filtered on total read volume, GC content, Q30 percentage and duplication rate. Later read alignment to human genome (hg19) was performed by BWA(9) (version 0.7.10). After sample coverage filtration, MuTect(10) (version 1.1.4) from GATK (version 4.0) pipeline identified single nucleotide variants (SNVs), small insertions and deletions (InDels) while segment-level somatic copy number alternations (SCNAs) were detected by GATK. Several rounds of filtrations on SNVs were conducted including (1) retaining variants with low frequency (<=0.01) in population from the 1000 Genomes Project (https://www.internationalgenome.org/), the Genome Aggregation Database (gnomAD) (https://gnomad.broadinstitute.org) and the Exome Aggregation Consortium (ExAC); (2) keeping mutations with number of supporting reads larger than 3; (3) keeping mutations with non-zero variant allele frequency (VAF) (>=0.01); (4) only functional alternations were preserved. Cancer-associated genes as well as cancer drive genes were collected for further comparisons. As for SCNAs, an in-house script(11) employed read coverages' statistical significance between tumor and normal tissues for focal level SCNA inference. Additionally, mutational spectrum and absolute contribution of COSMIC v3 SBS (single base substitution) mutational signatures were derived by MutationalPatterns(12) on unfiltered somatic mutations while Sigminer(13) quantified the absolute exposures of COSMIC v3 DBS (double base substitution) and ID (InDel) signatures.

Intratumor heterogeneity measurement and SNV/SCNA clonal architecture inference

We measured intratumor heterogeneity (ITH) of samples on both SNV and SCNA. For filtered SNVs, the mutant-allele tumor heterogeneity (MATH) score(14) was calculated using VAF values. The ABSOLUTE(15) tool further estimated the cancer ploidy, tumor purity, rescaled copy ratio and cancer cell fraction (CCF) combining SNV and SCNA data. The clonal architectures of SNVs were derived by the higher clonal mutation probability and the CCF upper 95% confidence interval larger than 1. As for SCNAs, copy neutral LOH (CNLOH) segments were initially discarded and clonal architectures were further annotated using allelic subclonal information from ABSOLUTE outputs.

DDR pathway gene analysis

DNA damage repair (DDR) related genes were collected from a previous publication(16) for the integrative analysis on SNV and SCNAs.

Inference of the clonal population structures

The populational structures on mutations were identified on filtered SNVs and annotated SCNAs by PyClone-VI(17). Later these clone clusters were visualized by ClonEvol(18).

Pathway annotation and Gene Oncology analysis

Enrichr(19) tool was used for pathway enrichment and Gene Oncology (GO) analysis. Enriched GO Biological Processes and Reactome(20) pathway entries were reported with P-values.

Public data curation for comparisons

For multi-omics data comparison, SNV, SCNA, transcriptomic and proteomics data were retrieved from multiple LUAD datasets. More specifically, SNV and SCNA data from four datasets (Lung-Broad(21), Lung-MSKCC, Lung-OncoSG(22) and TCGA-LUAD) were downloaded from cBioPortal for Cancer Genomics database(23) or UCSC Xena database(24) and only non-synonymous mutations were retained. Survival information was also downloaded if available. Additionally, transcriptomic data from Lung-OncoSG, TCGA-LUAD and one GEO dataset GSE148801(25) containing good-prognosis (e.g. LEP, ACI and PAP histological subtypes) and poor-prognosis (e.g. MIP and SOL) samples were collected while proteomics data from TCGA-LUAD were similarly curated. Only data from LEP and MIP subtypes were used for further comparisons.

Immune infiltration analysis

Immune infiltration analysis was conducted on curated transcriptomic samples for comparisons between LEP and MIP components. The RSEM(26) normalized expression as well as FPKM values were subjected to R package Immunedeconv(27) xCell(28) method for the derivation of density of immune cells in the tumor microenvironment (TME). Additionally, regarding the recognition, response and killing of cancer cells by immune system are in a step-wise manner, i.e. the cancer immunity cycle(29), activities of each step in the cycle were assessed by single sample Gene Set Enrichment Analysis (ssGSEA) from GSVA R package (http://bioconductor.org/packages/release/bioc/html/GSVA.html).

Statistical methods

Two-sided Mann-Whitney test was used for the evaluation of group-level differences between LEP and MIP components. As for multiple comparisons, P-values were adjusted by Benjamini-Hochberg method. Additionally, when the comparisons were conducted on categorical data, Fisher's exact test was utilized. As for the protein expression data, one-sided student's t-test was used in comparison. For all tests, a P-value < 0.05 was considered statistically significant. The Kaplan-Meier (K-M) survival curves were generated by survminer package(30) and the P-values were calculated by log-rank test.

Results

Clinicopathologic Characteristics of selected patients

Among the 286 patients passed initial filtering, 51 (17.8%) of them were LEP-predominant, 178 (62.2%) were ACI-predominant, 16 (5.6%) were PAP-predominant, 29 (10.1%) were MIP-predominant, and 12 (4.2%) were with SOL-predominant adenocarcinoma. The clinicopathologic characteristics of all patients were summarized in Supplementary Table 1. As shown in Supplementary Table 2, the 5-year DFS rates were 82.9%, 67.1%, 68.6%, 57.7% and 50.0% while the 5-year OS rates were 92.2%, 66.3%, 75.0%, 25.0% and 48.3% respectively for LEP-, ACI-, PAP-, MIP-, and SOL-predominant patients. Patients with MIP-predominant subtype exhibited a significantly worse DFS than those with other adenocarcinoma subtypes (P<0.05, Supplementary Figure 1A) and LEP and MIP subtypes exhibited largest DFS difference while SOL-predominant patients showed significantly worse OS (P<0.05, Supplementary Figure 1B).

Mutational landscape exhibits the involvement of distinct biological processes in LEP and MIP lesions

Among 8 micro-dissected samples, 6 cases passed quality control, but the quantity of LEP component in one case was not enough. 6 MIP and 5 LEP components were finally sequenced (Supplementary Table 3 and Figure 1A). A total of 2035 and 2757 SNVs and InDels were identified while 684/791 and 257/284 mutations were retained after quality and cancer-related gene filtrations (Supplementary Figure 2A and data quality metrics in Supplementary Table 4). Genes with top mutation frequency after quality filtering was shown in Figure 1B. *EGFR* was identified to be the most frequently mutated drive gene, in line with the finding that LEP and MIP components possess significantly higher *EGFR* mutation frequency(31). Besides, several cancer-associated genes including *TP53*, *TRIO*, *CEBPA*, *PCLO* and *PDE4DIP* were concomitantly mutated (Figure 1B),

denoting *p53*, WNT-beta-catenin signaling, *PI3K/AKT/mTOR* signaling and DNA repair pathways were affected. Interestingly, shared mutations were observed between paired LEP and MIP components from single patients (Figure 1B), raising the presumption that the paired LEP and MIP components could be homogenious. We also compared the mutation frequency of these genes with public cohorts, including Lung-Broad, Lung-OncoSG and TCGA-LUAD. Several cancer-associated genes including *EGFR*, *TP53*, *TRIO*, *PCLO* and *PDE4DIP* were found recurrently mutated (Supplementary Figure 2B).

Additionally, mutation signature analysis was conducted on unfiltered mutations separately for LEP and MIP components. The point substitution spectrum plot displayed insignificant difference between the two histological subtypes (Supplementary Figure 2C). Similarly, the SBS, InDel and DBS signatures mapped to COSMIC database (accessed in March 2021) were similar between the two subtypes (Figure 1C-D), denoting the histological differences between LEP and MIP components could be caused by alternations on specific key genes.

Of particular interest, mutated tumor suppressor genes (TSGs) were enriched in distinct pathways in LEP and MIP subtype (Figure 1E). TSGs from LEP components were enriched in DNA repair and *TP53*-related pathways while mutated TSGs in MIP components were found enriched in pathways associated with beta-catenin destruction complex, *AXIN* mutation and WNT signaling, which shared high consistency with previous report(4). When concerning the enriched pathways for mutated oncogenes, 7 out of 10 top enriched Reactome pathways from two groups were identical, which were mainly associated with *EGFR* and *PI3K* signaling (Supplementary Figure 2D). By further inspecting the mutated TSG pathway enrichment pattern in Lung-Broad (Supplementary Figure 3A-B), Lung-OncoSG (Supplementary Figure 3C-D) and TCGA-LUAD (Supplementary Figure 3E-F) cohorts, similar entries were identified in both LEP and MIP samples while *NOTCH1*-related pathways were addedly found in TCGA-LUAD MIP samples, which was unsurprising since the cross-talk between NOTCH and WNT pathways was previously unveiled(32). As for the oncogenes mutated in three public cohorts, shared terms were found between LEP and MIP components in Lung-Broad (Supplementary Figure 4A-B), Lung-OncoSG (Supplementary Figure 4C-D) as well as TCGA-LUAD (Supplementary Figure 4E-F) cohort but with lower overlapping proportion, again endorsed the possible existence of common mutational ancestor in the paired components.

Copy number alternation and clonality analysis uncovered distinct ITH characteristics in LEP and MIP subtypes

Through the segment level copy number alternation identification procedures, multiple amplified and deleted segments were detected (Supplementary Figure 5A). Chromosomes including 3,4,5,10,15,17 and 18 exhibited different copy number alternation patterns between two subtypes and the MIP subtype showed both higher chromosome level copy number variation (CNV) burden (Supplementary Figure 5B) and arm level CNV burden (Supplementary Figure 5C), which was consistent with previous report(33). To further pinpoint the recurrent SCNAs at focal level, we identified 1116 genes with somatic copy number alternations through statistical testing on read coverages from all samples, among which 159/80 genes were uniquely amplified/deleted in LEP component while 34/11 genes were uniquely amplified/deleted in MIP component. By further annotating the enriched pathways on these genes, 27 pathways were found overlapping between enrichment results of uniquely amplified genes in LEP and deleted genes in MIP, which could further be categorized into immune system, innate immune system, interleukin signaling, *SHC1* events, *ERK* activation and *FRS*-mediated signaling pathways (Figure 2A). When further inspecting the number of genes, pathways related to immune system and innate immune system got the highest gene number variated (37 genes amplified in LEP and 4 genes for deleted in MIP subtype), indicating MIP LUADs tend to have induced immunosurveillance escape. Additionally, two pathways were identical between the enrichment results of uniquely deleted genes in LEP and amplified genes in MIP (Figure 2B), which were associated with Homology Directed Repair (HDR) and mRNA fate regulation but the variated gene number was limited (6 genes for LEP- and 4 genes for MIP group).

The intratumor heterogeneity (ITH) can depict the genetic and epigenetic tumor inner diversity and was proven to be closely related to cancer progression, therapeutic resistance and recurrences. To compare the ITH of the two histological subtypes at both mutational and copy number level, we annotated the mutations/SCNAs with clonality. As shown in Figure 2C, MIP group has higher clonal tumor mutation burden (cTMB) and lower subclonal mutation proportion (Figure 2D), denoting the

lower mutation level ITH of the MIP subtype. The MATH score, which was widely used to measure the mutational ITH, exhibited similar trend (Supplementary Figure 5D). As for the copy number variations, MIP group possessed significantly higher proportion of subclonal SCNAs (Figure 2E) as well as subclonal genome fraction (Figure 2F). Interestingly, the frequency of clonal mutations in DNA Damage Response (DDR) and WNT pathway genes was higher in MIP subtype (Supplementary Figure 5E), which possibly partial give rise to the subclonal genome alternations on immune-related genes since the association between canonical WNT-beta-catenin signaling and carcinogenesis as well as immune suppression was clear(34). Indeed, 6 MIP components got higher percentage of subclonal SCNA (Figure 2G) as well as higher number of focal deletions (Figure 2H) on the genes related to the two immune pathways.

Evolutionary pattern exploration on the paired LEP and MIP components

To elaborate the possible evolutionary process between LEP and MIP subtypes, we delineated the phylogenetic trees for each patient based on mutations as well as focal level SCNAs. As shown in Figure 3, all 5 patients possessed truncal mutations between paired LEP and MIP components while no obvious bias on private mutation burden after truncal divergence was observed. Clonal mutations on cancer drivers including *EGFR*, *TP53* and *CEBPA* were identified and *EGFR* was the only gene coincident in 5 pairs, which confirmed the presence of ancestral mutations. Additionally, the driver mutations private to LEP were enriched in chromatin organization, *TP53*-related and DNA double strand repair pathways (Supplementary Figure 6A) while mutations private to MIP were enriched in cellular signaling and beta-catenin-related pathways (Supplementary Figure 6B). We further annotated the shared mutations in Figure 3 with clonality to explore the clonal-subclonal transitions between LEP and MIP subtype. For the genes possessing mutations with increased clonality in MIP, GO terms related to neurogenesis were found enriched (Supplementary Figure 6C), denoting the tumor-induced neurogenesis and nerve-cancer crosstalk may account for the aggressiveness of MIP subtype. Oppositely, genes with mutations switched from subclonal to clonal in LEP were associated with cell-cycle related GO biological processes (Supplementary Figure 6D). As for the truncal focal SCNAs, several driver genes including *CSMD3*, *SPTAN1*, *BCORL1*, *CAMTA1*, *GRIN2A*, *MED12* and *TRAF7* were concurrently amplified in the two subtypes (Figure 3), which was associated with developmental biology (R-HSA-1266738) and *EGFR*-related Reactome pathways (R-HSA-179812 and R-HSA-180336). Moreover, the deletion of *TP53*, *MUC4*, *ARID5B*, *ANK1*, *PTEN*, *SFPQ*, *FANCA*, *MAF* and *ZFHX3* were observed in the two subtypes. Interestingly, no driver gene showed concordant copy number variation in 5 pairs of samples, possibly due to the elevated SCNA-level ITH in MIP group. We also scrutinized the genes with shared copy number variation in the paired samples. As shown in Supplementary Figure 7A, most shared deletions were on immune-related genes, while signal transduction and *PI3K*/*AKT* pathways, which abnormality is highly associated with tumor progression and therapeutic resistance, were found uniformly amplified (Supplementary Figure 7B). To further derive the mutational transitions and evolutionary trajectory, we used PyClone-VI to infer the mutational populations and their evolution from paired components. As shown in Supplementary Figure 8A, numerous clone clusters were identified in 5 patients, which exhibited dynamic variant allele frequency (VAF) alternation. Clusters with drastically increased VAF in LEP subtype were mainly enriched in mRNA splicing pathways (Supplementary Figure 8B) while clusters with increased VAF in MIP subtype were associated with *ERBB2* functions (Supplementary Figure 8C). These data imply that LEP and MIP components from one patient were derived from same initiation cells and the pathway-specific mutations acquired after *EGFR* clonal mutation eventually shaped the subtype-specificity.

Group-wise comparison discovered possible novel therapeutic targets for MIP histological subtype

For the sake of the derivation of histological subtype-specific therapeutic targets, we next gathered SNV and SCNA data from the LEP/MIP (both LEP and MIP) groups and identified the genes with highest alternation frequency difference in all samples. As shown in Figure 4A, large mutation frequency difference was observed on 9 genes, with 3 genes specifically mutated in LEP group. Additionally, 5 genes were found with distinct copy number alternation pattern (Figure 4B), with 1 gene specifically amplified in LEP group. Similarly, mutation frequency on the 9 genes plus *EGFR* were compared between 4 public cohorts (Figure 4C) and *EGFR* was the only gene with significant mutational difference in non-east Asian public

cohorts. Moreover, all the 5 SCNA group-specific gene showed alternation frequency difference relatively in 3 non-east Asian cohorts (Figure 4D). Interestingly, the altered sample proportion or the alternation frequency for the MIP-group-amplified gene *PTP4A3*, *NAPRT* and *RECQL4* was highly similar in our samples (Figure 4B) and public cohorts (Lung-Broad, Lung-OncoSG and TCGA-LUAD, Figure 4D), implying the feasibility of their cooperative function through duplication in MIP component. Spearman's correlation coefficient (SCC) on SCNA pattern of our 11 samples confirmed the association of the co-amplified genes (Supplementary Figure 9A). Such strong association was also observed on LEP and MIP SCNA data from Lung-Broad (Supplementary Figure 9B), Lung-OncoSG (Supplementary Figure 9C, left) and TCGA-LUAD (Supplementary Figure 9D, left) cohorts. Concerning the fact that SCNA is highly related to the consequent gene expression alternation, we calculated the expressional SCC between the 5 group-specific genes on cohorts with available transcriptomic data. Unsurprisingly, when compared to all samples (Supplementary Figure 9C-D, middle), the expressional associations between *PTP4A3*, *NAPRT* and *RECQL4* transformed to a higher synergetic state for LEP/MIP samples in both Lung-OncoSG (Supplementary Figure 9C, right) and TCGA-LUAD (Supplementary Figure 9D, right) cohort. More explicitly, the correlation between SCNA and RNA expression was higher for the three genes in two public cohorts (Figure 4E) and *NAPRT* as well as *PTP4A3* exhibited significantly higher LEP/MIP group-specificity. As an exemplification, the correlation between *PTP4A3* gene's SCNA and RNA expression increased from 0.372 to 0.693 when narrowing to only LEP/MIP samples in TCGA-LUAD cohort (Supplementary Figure 9E-F). Complementarily, all three co-amplified genes were significantly overexpressed in tumor samples (Supplementary Figure 9G-I, left) and possessed higher expression levels in MIP subtype (Supplementary Figure 9G-I, right) in TCGA-LUAD cohort. Patients with elevated expressional level of the 3 co-amplified genes unsurprisingly got worse overall survival (OS) in Lung-OncoSG (Supplementary Figure 10A) and TCGA-LUAD (Supplementary Figure 10C) cohorts and the discrepancy exacerbated in LEP/MIP samples relatively in two cohorts (Supplementary Figure 10B,D). Complementally, patients with increased SCNA level on the co-amplified genes exhibited similar trend as the expressional stratification strategy, both for all LUAD samples (Supplementary Figure 10E) and LEP/MIP subset (Supplementary Figure 10F) in TCGA-LUAD cohort. To conclude, our comprehensive analyses identified *PTP4A3*, *NAPRT* and *RECQL4* as the possible novel therapeutic and diagnostic targets for MIP histological subtype, which were co-amplified and co-expressed specifically in LEP/MIP components and expressionally upregulated in MIP group.

Immune-related analyses uncovered elevated immunosuppression in MIP subtype

In connection with the 3 co-amplified genes and the possible exacerbated immunosuppression in MIP subtype, we quantified the immune infiltration in the tumor microenvironment (TME) of three public LUAD cohorts to investigate the possible roles of these genes. For Lung-OncoSG cohort, some differences were observed but the discrepancy did not reach statistical significance (Supplementary Figure 11A). We next correlated the SCNA level of 3 co-amplified genes with the immune cell infiltration levels. By only selecting correlations with P-value<0.01, entries including "Granulocyte-monocyte progenitor", "immune score" and "microenvironment score" were found negatively correlated (Supplementary Figure 11B), all indicated a reduced infiltration of general immune and stromal cells in MIP subtype. Similarly, when selecting correlations between RNA expression and immune infiltrations with P-value<0.01, *PTP4A3* was negatively correlated with CD4+ T memory cell infiltration while *NAPRT* expression was negatively associated with cancer-associated fibroblast (CAF) amount (Supplementary Figure 11C-D). As for the GEO dataset, significant between-group differences were observed on 7 cell types (Supplementary Figure 12A). Levels of hematopoietic stem cell (HSC) was elevated in relatively aggressive subtype (MIP/SOL) with adjusted P-value<0.01 while quantities of myeloid dendritic cell activated and mast cell decreased in aggressive subtypes. Naïve B cell, neutrophil and T cell CD4+ T-helper 1 cell infiltration were further found associated with *NAPRT* and *RECQL4* expressions (Supplementary Figure 12B). Similar with Lung-OncoSG cohort, the dissimilarities between LEP and MIP subtype were not statistically significant in TCGA-LUAD cohort (Supplementary Figure 13A). Similar cell types were also observed in the correlational analysis on SCNA-level (Supplementary Figure 13B) and RNA-level (Supplementary Figure 13C) data.

Moreover, inspection on the disparity of cancer immunity cycle activity was conducted on Lung-OncoSG, GEO and TCGA-LUAD cohorts. Activities of recurrent steps including release of cancer cell antigen, CD8+ T cell recruiting, dendritic cell

recruiting, macrophage recruiting, T-helper 17 (Th17) cell recruiting, T cell infiltration into tumors and killing of cancer cells were significantly higher in MIP subtype (Supplementary Figure 14A-C). Interestingly, the expressions of *PD-1* and *PD-L1* were generally higher in MIP group in Lung-OncoSG (Supplementary Figure 15A, D), GEO (Supplementary Figure 15B, E) and TCGA-LUAD (Supplementary Figure 15C, F) datasets. By further examining the differential expressed proteins between LEP and MIP subtype from TCGA-LUAD, the identified proteins with MIP-specific elevation (Supplementary Figure 15G) were significantly enriched in *PD-L1* and *PD-1* checkpoint pathway in cancer (Supplementary Figure 15H) and the protein-level expression of *PD-L1* was positively correlated with the transcriptional expression sum of co-amplified genes *PTP4A3*, *NAPRT* and *RECQL4* (Supplementary Figure 15I). Interestingly, the known associations between the co-amplified genes and chemotherapy resistance was previously reported(35). To summarize briefly, our analyses confirmed the immunosuppression prevalence in MIP subtype and provided therapeutic suggestions for MIP-predominant LUAD patients. Shaped by the higher T cell infiltration level, possible chemotherapy resistance and immunosuppression, immune checkpoint inhibitor treatment could maximize the therapeutic benefit for MIP-predominant LUAD patients.

Discussion

LUAD exhibited high inner heterogeneity, including 5 subtypes: lepidic, acinar, papillary, micropapillary and solid. Consistent with other studies, we confirmed the survival disparity between LUAD subtypes. By performing WES on micro-dissected LUAD tissue samples of MIP and LEP components, we explored the genetic feature related to LEP/MIP growth pattern and evolutional connection between LUAD subtypes. Our results revealed that LEP and MIP subtypes could be derived from same initiation cells with *EGFR* mutation and the ultimate histological dissimilitude was shaped by the pathway-specific mutations acquired along evolution. Previously Wang et al. revealed trunk mutations of common oncogenes like *EGFR*, *TP53* played a dominant role in early invasive LUAD(36) while Xin et al. observed truncal *EGFR* and *KRAS* mutations between preneoplasia and invasive LUAD in multi-region multifocal pulmonary nodules from the same patients(37). Our results showed *EGFR* trunk mutation arose between pre-invasion and invasion LUAD components and LEP/MIP components were evolved by branched evolution model.

Through comprehensively comparisons on genetic alternations, biological characteristics of the two LUAD subtypes were further elucidated. As for mutational comparisons, TSG mutations in LEP were associated with DNA repair and *TP53* regulation while genes related to WNT signaling and beta-catenin destruction complex got both higher mutational frequency and clonality. Driver mutations private to MIP were also enriched in cellular signaling and beta-catenin-related pathways, while genes possessed lower mutational heterogeneity in MIP were associated with neurogenesis and *ERBB2* functions. Aberrant WNT signaling pathway activation caused by gene mutations of intracellular components is associated with a higher rate of recurrence in early-stage NSCLC. On the other side, being the critical downstream effector in canonical WNT pathway, excessive intracellular beta-catenin promotes lung cancer aggression. Liang et. al further confirmed the intracellular beta-catenin expression in MIP-predominant LUAD was higher than LEP-predominant LUAD(4). Besides, neurogenesis induced by tumor shapes a immunosuppressive microenvironment(38). Although cancer-related neurogenesis has been considered to be associated with solid tumor metastasis, its role in LUAD remains poorly understood. Our results suggested the inner association between MIP aggressive phenotype and neurogenesis. The activation of well-known proto-oncogene *ERBB2* signaling was associated with poor outcomes in NSCLC(39), again coincides with MIP characteristics. Additionally, copy number of genes related to immune system, innate immune system, interleukin signaling, SHC1 events, ERK activation and FRS-mediated signaling were also found decreased in MIP. With highest proportion of immune genes affected, the immunosuppression status in MIP subtype was confirmed. Apart from specified genomic alternations, ITH provides crucial information for drug responsiveness and clinical prognosis. Discordance between SNV and SCNA ITH was particularly observed in MIP subtype. Subclonal genetic instability possibly facilitated MIP neoplastic cell proliferation(40) and the clonal mutations on key MIP-specific pathways contributed to its aggressive behavior.

The current diagnosis of LUAD growth patterns was based on morphologic assessment, which mainly relies on the subjective experience of pathologists. We discovered three potential diagnostic and therapeutic target genes with co-

amplification tendency in MIP, both in our discovery cohort and three public validation cohorts. Previous studies proved the knockdown of *PTP4A3* inhibited cell migration and invasion of lung cancer cell lines(41). It also induced microvascular and lymphatic vessel formation by fascinating VEGF and VEGF-C expression in lung cancer tissues, which was in accordance with the clinical observations that MIP component in LUAD increases the risk of distant and lymph node metastasis. Previous study found the loss of *NAPRT* promoted the Epithelial-Mesenchymal Transition (EMT) by stabilizing beta-catenin(42). The elevated expression of *NAPRT* conceivably was associated with the disruption of catenin-cadherin complex in MIP. Besides, *RECQL4* could coordinate and regulate cell proliferation and cell cycle progression by protecting chromosome stability(43) and its protein expression was remarkably higher in LUAD(44). These biological mechanisms of the three genes further verified our discoveries.

Surgical resection, chemotherapy, and targeted therapy are common therapeutic options for MIP-predominant LUAD but limitations still exist respectively. Regarding the increasing enthusiasm on lung cancer immunotherapy, we assessed and compared the TME landscape between MIP and LEP components. The three co-amplified genes showed associations with cell infiltration levels of general immune/stromal cells, CD4+ T memory cells and CAF. Reduced infiltration of general immune and stromal cells in MIP subtype indicated its lower systemic immunity status, lower amount of CD4+ T memory cell was associated with worse survival in other tumor type(45) while CAFs were connected to decreased number of CD8+ T cells and possible immunotherapy resistance(46). Additionally, MIP/SOL subtype got elevated infiltration of HSC, which was associated with immunosuppressive phenotypes and elevated PD-L1 expression in glioblastoma(47). Mast cell infiltration level decreased in MIP/SOL samples and its depletion improved the efficacy of anti-PD-1 therapy in melanoma model(48). As for the cancer immunity cycle stepwise activity, steps including CD8+ T cell recruiting, T-helper 17 (Th17) cell recruiting, T cell infiltration into tumors and killing of cancer cells were significantly higher in MIP samples. In addition to the TME analyses, proteomics analysis further confirmed the immunosuppression status specificity in MIP subtype. Unsurprisingly, the PD-L1 protein expression correlated with the three co-amplified genes and their known associations with chemotherapy resistance was revealed previously. In general, our work revealed the comprehensive TME situation of LEP/MIP components and emphasized the suitableness of MIP-predominant patients on anti-PD-1/anti-PD-L1 immunotherapies.

Our study exists several limitations. Firstly, the cohort only included 5 pairs of LEP/MIP components detached from 5 LUAD patients and a total of 11 samples. Further studies with larger amount of patient involvement can better decipher the evolutionary trajectory between LUAD histological subtypes and identify subtype-specific genetic changes. Moreover, the three potential diagnostic and therapeutic target genes with co-amplification or co-expression tendency in MIP should be further experimentally validated, especially on their protein expression status. Additionally, we portrayed the TME heterogeneity using bulk RNA-seq data. With the recent maturation of multiple advanced techniques, utilizing methods including single-cell RNA-seq, spatial transcriptomics and multiplexed immunohistochemistry could better dissect the TME in LUAD. Lastly, our analyses only focused on MIP and LEP subtypes. A more integrated study incorporating other LUAD histologic subtypes could better decode the disease.

To conclude, by selecting patients with both MIP and LEP components in one lesion for WES, we identified subtype-specific genetic differences responsible for variated prognosis. We also revealed the evolution trajectory of MIP subtype. The subtype-specificity was possibly shaped by pathway-specific mutations acquired after *EGFR* clonal mutation. Besides, by utilizing data from multiple cohorts, genes specifically altered in MIP subtype were identified as novel diagnostic and therapeutic targets. Tumor microenvironment and proteomics analyses also revealed the immunosuppression prevalence in MIP. Immune checkpoint inhibitor treatments like anti-PD-1/anti-PD-L1 possibly could maximize the therapeutic benefit for MIP-predominant LUAD patients.

Abbreviations

NSCLC	non-small-cell lung cancer	SCNAs	somatic copy number alternations
WHO	world health organization	VAF	variant allele frequency
LUAD	lung adenocarcinoma	SBS	single base substitution
MIP	micropapillary growth pattern	DBS	double base substitution
LEP	lepidic growth pattern	ID	InDel
ACI	acinar growth pattern	ITH	intratumor heterogeneity
PAP	papillary growth pattern	MATH	mutant-allele tumor heterogeneity
SOL	solid growth pattern	CCF	cancer cell fraction
DFS	disease-free survival	CNLOH	copy neutral LOH
OS	overall survival	DDR	DNA damage repair
TMB	tumor mutation burden	GO	Gene Oncology
TKI	tyrosine kinase inhibitor	K-M	Kaplan-Meier
PD-L1	programmed death-ligand 1	TSGs	tumor suppressor genes
PD-1	programmed cell death protein 1	CNV	copy number variation
TME	tumor microenvironment	HDR	Homology Directed Repair
WES	whole-exome sequencing	SCC	Spearman's correlation coefficient
FFPE	formalin-fixed paraffin-embedded	CAF	cancer-associated fibroblast
H&E	hematoxylin and eosin	HSC	hematopoietic stem cell
SNVs	single nucleotide variants	Th17	T-helper 17
InDels	insertions and deletions	EMT	Epithelial-Mesenchymal Transition

Declarations

Availability of data and materials

The sequencing data on six LUAD patients are available from the corresponding author upon reasonable request. Besides, the directories of the public datasets used for validation are listed below:

Lung-Broad: https://www.cbioportal.org/study/summary?id=luad_broad;

Lung-MSKCC: https://www.cbioportal.org/study/summary?id=luad_mskcc_2020;

Lung-OncoSG: https://www.cbioportal.org/study/summary?id=luad_oncosg_2020;

TCGA-LUAD: https://xenabrowser.net/datapages/?cohort=TCGA%20Lung%20Adenocarcinoma%20(LUAD)&removeHub=https%3A%2F%2Fxena.treehouse.gi.ucsc.edu%3A443;

GSE148801: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE148801.

Acknowledgements

Not applicable.

Funding

This work was supported by National Natural Science Foundation of China (grant number 81772484 to CLW, grant number 82173038 to DSY)

Author Contributions

Changli Wang, Bin Zhang and Dongsheng Yue contributed to the study design. Leina Sun, Zhenfa Zhang, Lianmin Zhang, Yansong Huo, Wuhao Huang, Xiaoyan Sun, Zhe Tang, Yingnan Feng and Huilan Mo contributed to data collection, Jie Yuan, Hua Zhang, Zuoquan Yang, Chao Zhang and Zicheng Yu performed statistical analysis, interpretation, and drafted the manuscript. Changli Wang and Bin Zhang revised the manuscript. All authors contributed to critical revision of the final manuscript. Changli Wang and Dongsheng Yue provided financial support and study supervision. All authors read and approved the final manuscript.

Ethics declarations

Ethics approval and consent to participate

This study was ratified by the Ethics Committee of Tianjin Cancer Hospitoal. And this study does not involve any animal experiments. All the public data used in this article were from open assess databases, and no permission was required.

Consent for publication

All authors have approved the final manuscript for publication.

Competing interests

The authors declare that they have no competing interests.

References

1. Bray F, Ferlay J, Laversanne M, Brewster DH, Gombe Mbalawa C, Kohler B, et al. Cancer Incidence in Five Continents: Inclusion criteria, highlights from Volume X and the global status of cancer registration. Int J Cancer. 2015;137(9):2060-71.
2. Hung JJ, Yeh YC, Jeng WJ, Wu YC, Chou TY, Hsu WH. Factors predicting occult lymph node metastasis in completely resected lung adenocarcinoma of 3 cm or smaller. Eur J Cardiothorac Surg. 2016;50(2):329-36.
3. Makinen JM, Laitakari K, Johnson S, Makitaro R, Bloigu R, Lappi-Blanco E, et al. Nonpredominant lepidic pattern correlates with better outcome in invasive lung adenocarcinoma. Lung Cancer. 2015;90(3):568-74.
4. Zhu L, Yang S, Zheng L, Zhang G, Cheng G. WNT/beta-catenin pathway activation via Wnt1 overexpression and Axin1 downregulation correlates with cadherin-catenin complex disruption and increased lymph node involvement in micropapillary-predominant lung adenocarcinoma. J Thorac Dis. 2020;12(10):5906-15.
5. Warth A, Penzel R, Lindenmaier H, Brandt R, Stenzinger A, Herpel E, et al. EGFR, KRAS, BRAF and ALK gene alterations in lung adenocarcinomas: patient outcome, interplay with morphology and immunophenotype. Eur Respir J. 2014;43(3):872-83.
6. Tsao MS, Marguet S, Le Teuff G, Lantuejoul S, Shepherd FA, Seymour L, et al. Subtype Classification of Lung Adenocarcinoma Predicts Benefit From Adjuvant Chemotherapy in Patients Undergoing Complete Resection. J Clin Oncol. 2015;33(30):3439-46.
7. Ng Kee Kwong F, Laggner U, McKinney O, Croud J, Rice A, Nicholson AG. Expression of PD-L1 correlates with pleomorphic morphology and histological patterns of non-small-cell lung carcinomas. Histopathology.

2018;72(6):1024-32.
8. Zhang S, Xu Y, Zhao P, Bao H, Wang X, Liu R, et al. Integrated Analysis of Genomic and Immunological Features in Lung Adenocarcinoma With Micropapillary Component. Front Oncol. 2021;11:652193.
9. Li HJapa. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. 2013.
10. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. 2013;31(3):213-9.
11. McGranahan N, Rosenthal R, Hiley CT, Rowan AJ, Watkins TB, Wilson GA, et al. Allele-specific HLA loss and immune escape in lung cancer evolution. 2017;171(6):1259-71. e11.
12. Blokzijl F, Janssen R, van Boxtel R, Cuppen EJGm. MutationalPatterns: comprehensive genome-wide analysis of mutational processes. 2018;10(1):1-11.
13. Wang S, Tao Z, Wu T, Liu X-SJB. Sigflow: an automated and comprehensive pipeline for cancer genome mutational signature analysis. 2021;37(11):1590-2.
14. Mroz EA, Rocco JWJOo. MATH, a novel measure of intratumor genetic heterogeneity, is high in poor-outcome classes of head and neck squamous cell carcinoma. 2013;49(3):211-5.
15. Carter SL, Cibulskis K, Helman E, McKenna A, Shen H, Zack T, et al. Absolute quantification of somatic DNA alterations in human cancer. 2012;30(5):413-21.
16. Knijnenburg TA, Wang L, Zimmermann MT, Chambwe N, Gao GF, Cherniack AD, et al. Genomic and molecular landscape of DNA damage repair deficiency across The Cancer Genome Atlas. 2018;23(1):239-54. e6.
17. Gillis S, Roth AJBb. PyClone-VI: scalable inference of clonal population structures using whole genome data. 2020;21(1):1-16.
18. Dang H, White B, Foltz S, Miller C, Luo J, Fields R, et al. ClonEvol: clonal ordering and visualization in cancer sequencing. 2017;28(12):3076-82.
19. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. 2016;44(W1):W90-W7.
20. Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, et al. The reactome pathway knowledgebase. 2018;46(D1):D649-D55.
21. Imielinski M, Berger AH, Hammerman PS, Hernandez B, Pugh TJ, Hodis E, et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. 2012;150(6):1107-20.
22. Chen J, Yang H, Teo ASM, Amer LB, Sherbaf FG, Tan CQ, et al. Genomic landscape of lung adenocarcinoma in East Asians. 2020;52(2):177-86.
23. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. 2013;6(269):pl1-pl.
24. Goldman M, Craft B, Brooks A, Zhu J, Haussler DJb. The UCSC Xena Platform for cancer genomics data visualization and interpretation. 2018:326470.
25. Tone M, Tahara S, Nojima S, Motooka D, Okuzaki D, Morii EJCs. HTR3A is correlated with unfavorable histology and promotes proliferation through ERK phosphorylation in lung adenocarcinoma. 2020;111(10):3953.
26. Li B, Dewey CNJBb. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. 2011;12(1):1-16.
27. Sturm G, Finotello F, Petitprez F, Zhang JD, Baumbach J, Fridman WH, et al. Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. 2019;35(14):i436-i45.
28. Aran D, Hu Z, Butte AJJGb. xCell: digitally portraying the tissue cellular heterogeneity landscape. 2017;18(1):1-14.
29. Chen DS, Mellman IJi. Oncology meets immunology: the cancer-immunity cycle. 2013;39(1):1-10.
30. Kassambara A, Kosinski M, Biecek P, Fabian SJRpv. survminer: Drawing Survival Curves using 'ggplot2'. 2017;1.

31. Li H, Pan Y, Li Y, Li C, Wang R, Hu H, et al. Frequency of well-identified oncogenic driver mutations in lung adenocarcinoma of smokers varies with histological subtypes and graduated smoking dose. Lung Cancer. 2013;79(1):8-13.
32. Krishnamurthy N, Kurzrock RJCtr. Targeting the Wnt/beta-catenin pathway in cancer: Update on effectors and inhibitors. 2018;62:50-60.
33. Caso R, Sanchez-Vega F, Tan KS, Mastrogiacomo B, Zhou J, Jones GD, et al. The underlying tumor genomics of predominant histologic subtypes in lung adenocarcinoma. 2020;15(12):1844-56.
34. Chehrazi-Raffle A, Dorff T, Pal S, Lyou Y. Wnt/β-Catenin Signaling and Immunotherapy Resistance: Lessons for the Treatment of Urothelial Carcinoma. Cancers 2021, 13, 889. s Note: MDPI stays neutral with regard to jurisdictional claims in published …; 2021.
35. Li X-q, Lei J, Mao L-h, Wang Q-l, Xu F, Ran T, et al. NAMPT and NAPRT, key enzymes in NAD salvage synthesis pathway, are of negative prognostic value in colorectal cancer. 2019;9:736.
36. Wang S, Du M, Zhang J, Xu W, Yuan Q, Li M, et al. Tumor evolutionary trajectories during the acquisition of invasiveness in early stage lung adenocarcinoma. Nat Commun. 2020;11(1):6083.
37. Hu X, Fujimoto J, Ying L, Fukuoka J, Ashizawa K, Sun W, et al. Multi-region exome sequencing reveals genomic evolution from preneoplasia to lung adenocarcinoma. 2019;10(1):1-10.
38. Cervantes-Villagrana RD, Albores-García D, Cervantes-Villagrana AR, García-Acevez SJJSt, therapy t. Tumor-induced neurogenesis and immune evasion as targets of innovative anti-cancer therapies. 2020;5(1):1-23.
39. Riudavets M, Sullivan I, Abdayem P, Planchard DJEo. Targeting HER2 in non-small-cell lung cancer (NSCLC): a glimpse of hope? An updated review on therapeutic strategies in NSCLC harbouring HER2 alterations. 2021;6(5):100260.
40. Dentro SC, Leshchiner I, Haase K, Tarabichi M, Wintersinger J, Deshwar AG, et al. Characterizing genetic intra-tumor heterogeneity across 2,658 human cancer genomes. 2021;184(8):2239-54. e39.
41. Jian M, Nan L, Guocheng J, Qingfu Z, Xueshan Q, Enhua WJTJ. Downregulating PRL-3 inhibit migration and invasion of lung cancer cell via RhoA and mDia1. 2012;98(3):370-6.
42. Lee J, Kim H, Lee JE, Shin S-J, Oh S, Kwon G, et al. Selective cytotoxicity of the NAMPT inhibitor FK866 toward gastric cancer cells with markers of the epithelial-mesenchymal transition, due to loss of NAPRT. 2018;155(3):799-814. e13.
43. Fang H, Niu K, Mo D, Zhu Y, Tan Q, Wei D, et al. RecQL4-Aurora B kinase axis is essential for cellular proliferation, cell cycle progression, and mitotic integrity. 2018;7(9):1-12.
44. Jiang W, Xu J, Liao Z, Li G, Zhang C, Feng YJFic, et al. Prognostic Signature for Lung Adenocarcinoma Patients Based on Cell-Cycle-Related Genes. 2021;9:635.
45. Ning Z-K, Hu C-G, Huang C, Liu J, Zhou T-C, Zong ZJFio. Molecular subtypes and CD4+ memory T cell-based signature associated with clinical outcomes in gastric cancer. 2021;10:3188.
46. Min K-W, Kim D-H, Noh Y-K, Son BK, Kwon MJ, Moon J-YJSr. Cancer-associated fibroblasts are associated with poor prognosis in solid type of lung adenocarcinoma in a machine learning analysis. 2021;11(1):1-9.
47. Lu I-N, Dobersalske C, Rauschenbach L, Teuber-Hanselmann S, Steinbach A, Ullrich V, et al. Tumor-associated hematopoietic stem and progenitor cells positively linked to glioblastoma progression. 2021;12(1):1-16.
48. Somasundaram R, Connelly T, Choi R, Choi H, Samarkina A, Li L, et al. Tumor-infiltrating mast cells are associated with resistance to anti-PD-1 therapy. 2021;12(1):1-14.

Figures

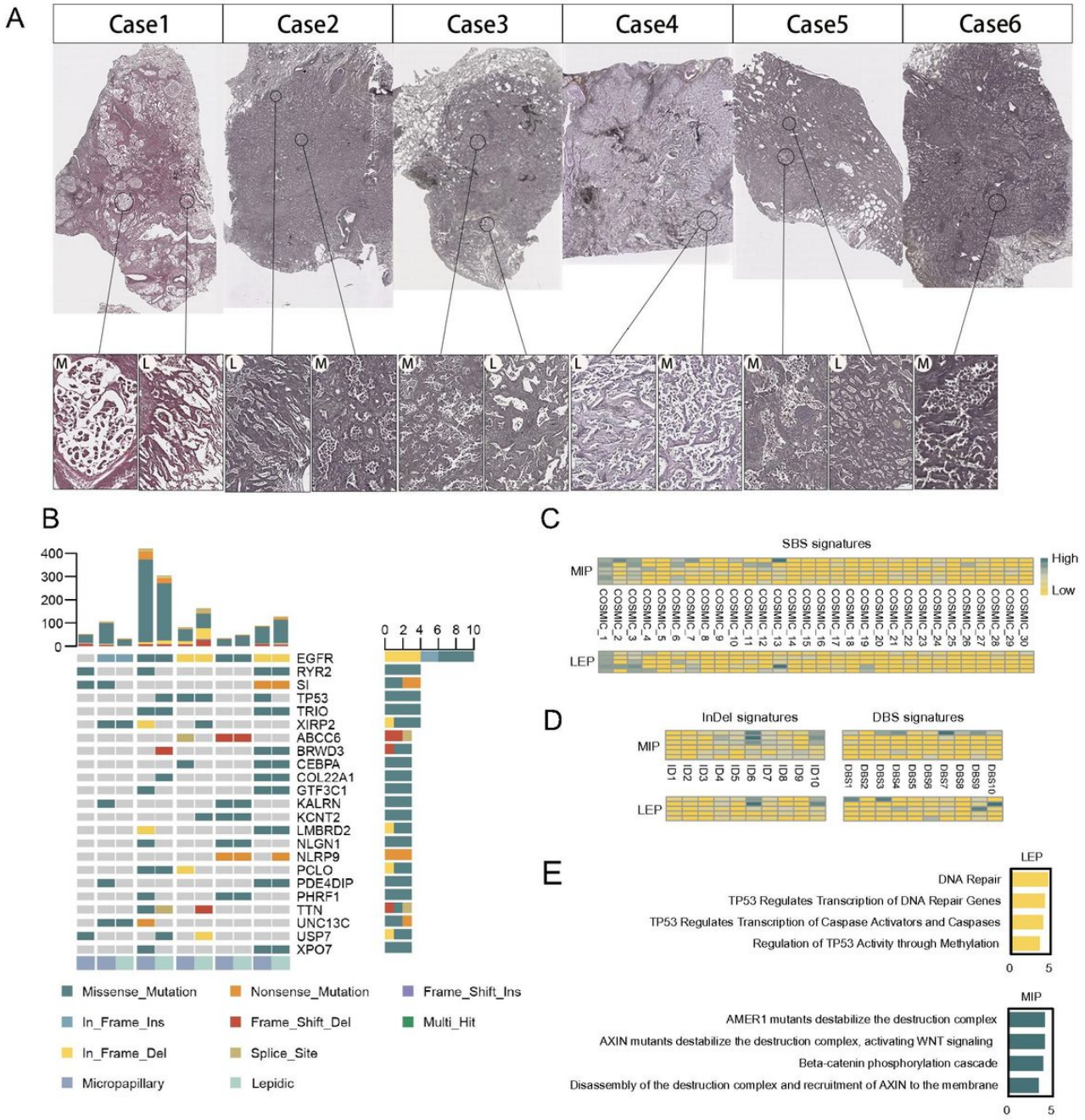

Figure 1

The H&E stained slides and the mutational landscape of MIP and LEP subtypes. (A) Scanner view (magnification 20x) of the H&E-stained MIP (M) and LEP (L) subtypes from six patients. (B) Genes with top mutation frequency in samples. Cancer-associated genes were marked with red circles. (C) Quantifications on the SBS signatures in two histological subtypes. (D) Quantifications on the InDel and DBS signatures in subtypes. (E) Pathways enriched in mutated tumor suppressor genes (TSGs) for MIP and LEP subtypes.

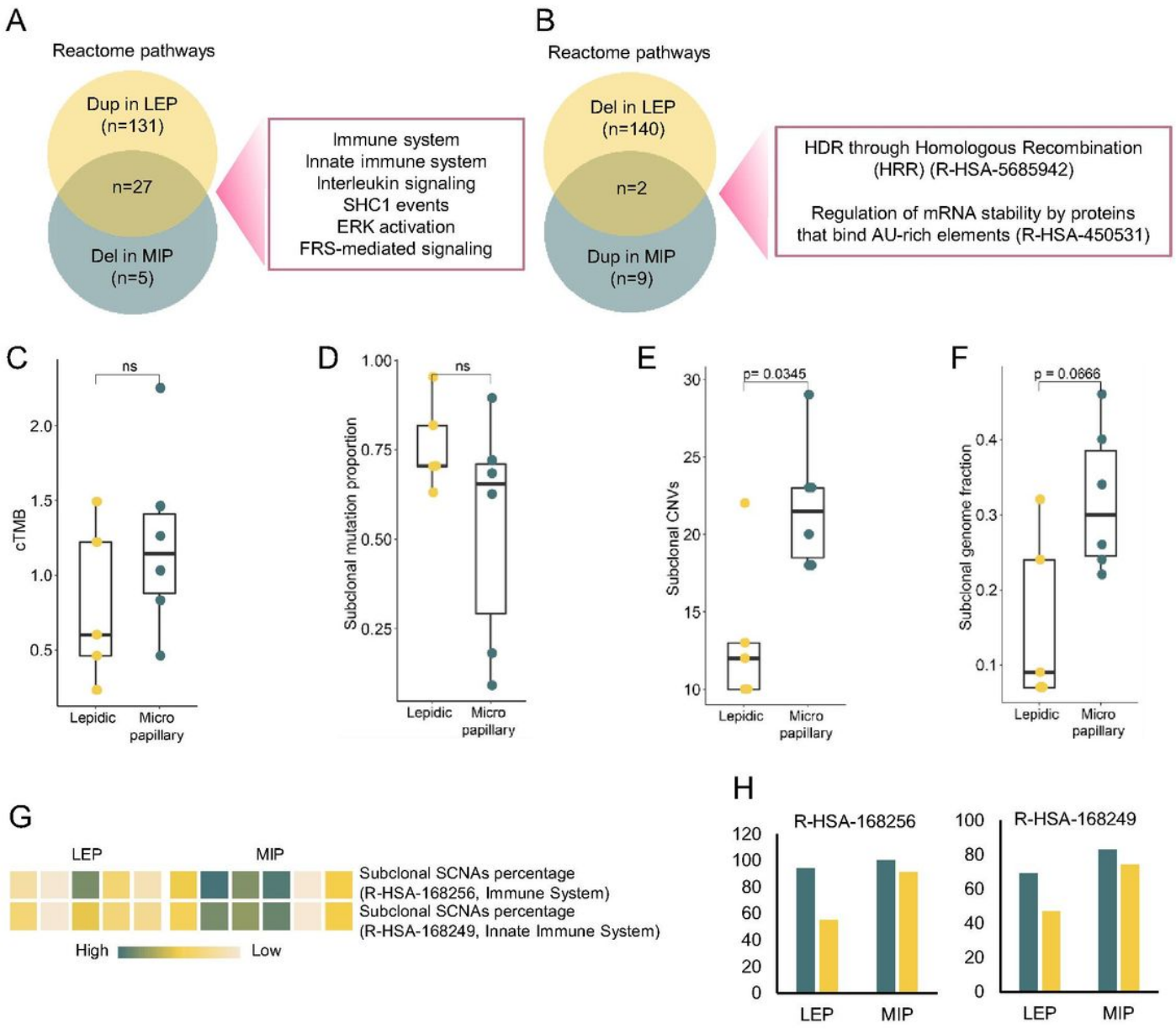

Figure 2

Multi-perspective investigation on intratumor heterogeneity (ITH) difference between two subtypes. (A) Intersection on the enriched pathways in genes uniquely amplified in LEP and deleted in MIP. (B) Intersection on the enriched pathways in genes uniquely deleted in LEP and amplified in MIP. (C-F) Clonal tumor mutation burden (cTMB), subclonal mutation proportion, subclonal CNV and subclonal genome fraction distribution in two subtypes. (G) Subclonal SCNA percentage of two Reactome immune pathways in sequenced samples. (H) Focal alternation number on the genes in two immune pathways. P-values on the alternation discrepancy between subtypes were calculated by Fisher's exact test.

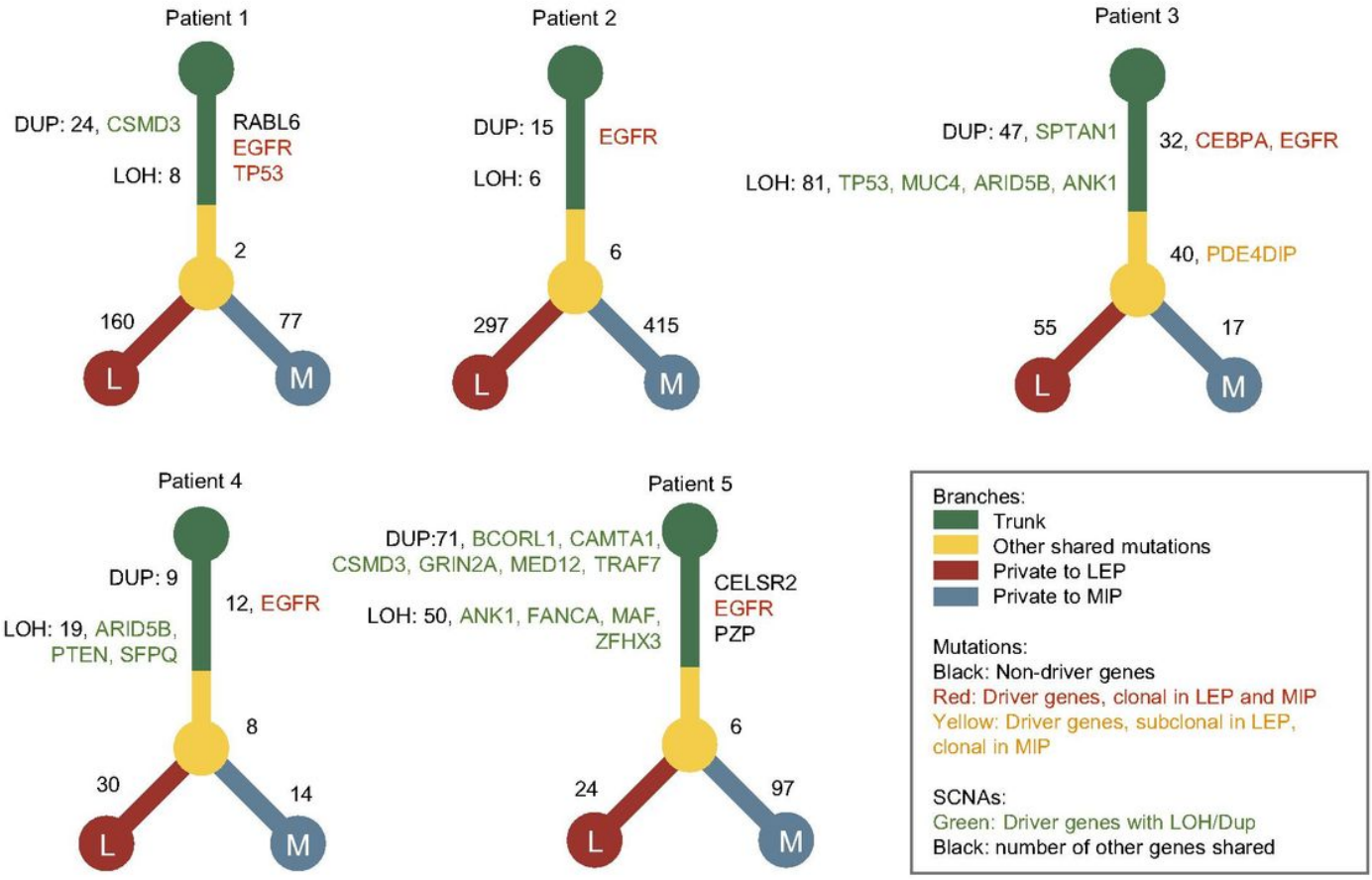

Figure 3

The phylogenetic trees constructed for patients with concomitant MIP and LEP components. Driver genes with mutations and focal CNV were marked by different colors. Number of mutations and focal CNVs were labelled beside the tree trunk while mutation number were labeled beside tree branches.

A

	Number of samples		Number of mutations		
	LEP	MIP	LEP	MIP	Type
C10orf71	1	0	3	0	Mut
SLC32A1	1	0	4	0	Mut
DISC1	1	0	5	0	Mut
AHCTF1	0	1	0	3	Mut
PHRF1	1	2	1	3	Mut
PLEC	0	1	0	3	Mut
RYR2	1	3	1	3	Mut
SI	1	3	0	3	Mut
SYNE2	1	2	1	4	Mut

B

	Number of samples		
	LEP	MIP	Type
RCSD1	3	0	Dup
PTP4A3	1	5	Dup
EZR	2	5	Del
NAPRT	2	5	Dup
RECQL4	2	5	Dup

C

	Lung-Broad				Lung-MSKCC				Lung-OncoSG				TCGA-LUAD			
	Number of samples															
	LEP (n=13)	MIP (n=17)	Type	Adj. P-value	LEP (n=88)	MIP or Solid (n=105)	Type	P-value	LEP (n=10)	MIP (n=4)	Type	Adj. P-value	LEP (n=12)	MIP (n=23)	Type	Adj. P-value
EGFR	4	2	Mut	0.000518	38	19	Mut	7.272e-05	6	3	Mut	0.525	2	4	Mut	0.002464
SLC32A1	0	1	Mut	0.626									1	1	Mut	1
DISC1													0	1	Mut	1
AHCTF1	0	1	Mut	0.626									2	1	Mut	1
PHRF1	0	1	Mut	0.13904									0	1	Mut	0.340335
PLEC	0	3	Mut	1									1	2	Mut	1
RYR2	1	7	Mut	0.801257					2	1	Mut	0.656	2	7	Mut	1
SI	1	5	Mut	0.626					2	0	Mut	0.525	1	4	Mut	1
SYNE2	0	1	Mut	0.13904									1	4	Mut	1

D

	Lung-Broad				Lung-MSKCC				Lung-OncoSG				TCGA-LUAD			
	Number of samples															
	LEP (n=13)	MIP (n=17)	Type	Adj. P-value	LEP (n=88)	MIP or Solid (n=105)	Type	P-value	LEP (n=10)	MIP (n=4)	Type	Adj. P-value	LEP (n=12)	MIP (n=23)	Type	Adj. P-value
RCSD1	4	4	Dup	1					6	3	Dup	0.5535	10	20	Dup	0.002777
PTP4A3	3	1	Dup	0.0502					5	2	Dup	1	6	10	Dup	1
EZR	0	1	Del	0.0004212					2	2	Del	0.58	8	13	Del	1
NAPRT	3	1	Dup	0.012328					5	2	Dup	1	6	10	Dup	1
RECQL4					2	1	Dup	1.427e-08	5	2	Dup	1	6	10	Dup	1

E

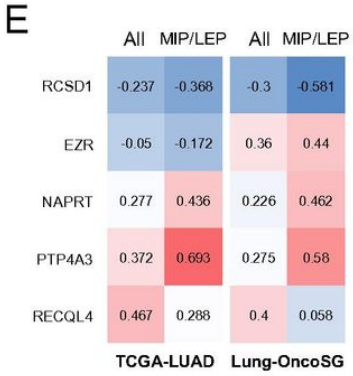

Figure 4

Identification of genes with subtype-specific alternations. (A) Genes with mutational specificity in LEP or MIP components. (B) Genes with focal SCNA specificity in LEP or MIP. (C) Mutation frequency of the subtype-specific mutated genes and EGFR in 4 public cohorts. (D) Focal SCNA frequency of the subtype-specific altered genes in 4 public cohorts. (E) Spearman's correlation coefficient (SCC) calculated using focal SCNA and expression values of five subtype-specific genes in two public LUAD cohorts.

Supplementary Files

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

- Supplementfigure1.pdf
- Supplementfigure2.pdf

- Supplementfigure3.pdf
- Supplementfigure4.pdf
- Supplementfigure5.pdf
- Supplementfigure6.pdf
- Supplementfigure7.pdf
- Supplementfigure8.pdf
- Supplementfigure9.pdf
- Supplementfigure10.pdf
- Supplementfigure11.pdf
- Supplementfigure12.pdf
- Supplementfigure13.pdf
- Supplementfigure14.pdf
- Supplementfigure15.pdf
- SupplementaryTable1.xlsx
- SupplementaryTable2.xlsx
- SupplementaryTable3.xlsx
- SupplementaryTable4.xlsx
- Supplementaryfigurelegends.docx