KJTCVS
The Korean Journal of Thoracic and Cardiovascular Surgery
https://doi.org/10.5090/kjtcs.20.013
pISSN: 2233-601X eISSN: 2093-6516
Korean J Thorac Cardiovasc Surg. 2020;53(5):250-257

Basic Research

Genetic Features of Lung Adenocarcinoma with Ground-Glass Opacity: What Causes the Invasiveness of Lung Adenocarcinoma?

Dohun Kim, M.D., Ph.D.[1], Jong-Young Lee, Ph.D.[2], Jin Young Yoo, M.D.[3], Jun Yeun Cho, M.D.[4]

[1]Department of Thoracic and Cardiovascular Surgery, Chungbuk National University Hospital, Chungbuk National University College of Medicine, Cheongju; [2]Institute of Genomic Health, Oneomics Co. Ltd., Seoul; [3]Department of Radiology and [4]Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Chungbuk National University Hospital, Cheongju, Korea

ARTICLE INFO
Received March 20, 2020
Revised May 13, 2020
Accepted May 25, 2020

Corresponding author
Dohun Kim
Tel 82-43-269-6061
Fax 82-43-269-6969
E-mail mwille@chungbuk.ac.kr
ORCID
https://orcid.org/0000-0001-8304-0232

Jun Yeun Cho
Tel 82-43-269-6029
Fax 82-43-269-7614
E-mail kaist2115@cbnuh.or.kr
ORCID
https://orcid.org/0000-0003-4270-413X

Background: Lung adenocarcinoma (LUAD) with ground-glass opacity (GGO) can become aggravated, but the reasons for this aggravation are not fully understood. The goal of this study was to analyze the genetic features and causes of progression of GGO LUAD.
Methods: LUAD tumor samples and normal tissues were analyzed using an Illumina HiSeq 4000 system. After the tumor mutational burden (TMB) was calculated, the identified mutations were classified as those found only in GGO LUAD, those present only in non-GGO LUAD, and those common to both tissue types. Ten high-frequency genes were selected from each domain, after which protein interaction network analysis was conducted.
Results: Overall, 227 mutations in GGO LUAD, 212 in non-GGO LUAD, and 48 that were common to both tumor types were found. The TMB was 8.8 in GGO and 7.8 in non-GGO samples. In GGO LUAD, mutations of *FCGBP* and *SFTPA1* were identified. *FOXQ1*, *IRF5*, and *MAGEC1* mutations were common to both types, and *CDC27* and *NOTCH4* mutations were identified in the non-GGO LUAD. Protein interaction network analysis indicated that *IRF5* (common to both tissue types) and *CDC27* (found in the non-GGO LUAD) had significant biological functions related to the cell cycle and proliferation.
Conclusion: In conclusion, GGO LUAD exhibited a higher TMB than non-GGO LUAD. No clinically meaningful mutations were found to be specific to GGO LUAD, but mutations involved in the epithelial-mesenchymal transition or cell cycle were found in both tumor types and in non-GGO tissue alone. These findings could explain the non-invasiveness of GGO-type LUAD.

Keywords: Adenocarcinoma of lung, Mutation, High-throughput nucleotide sequencing, DNA, DNA mutational analysis

Introduction

Advances in imaging technology have increased the probability of the early detection of lung cancer, especially lung adenocarcinoma (LUAD) with ground-glass opacity (GGO) [1,2]. This type of adenocarcinoma has an excellent survival rate compared to non-GGO adenocarcinomas [3,4]. GGO-type LUAD can be divided into pure and mixed GGO. Pure GGO (which lacks regions of consolidation) is likely to be adenocarcinoma *in situ* (AIS) or minimally invasive adenocarcinoma (MIA), while mixed GGO (which exhibits areas of consolidation) is likely to be MIA, lepidic predominant invasive adenocarcinoma, or a more invasive type of adenocarcinoma. Therefore, pure GGO LUAD can be managed through sublobar resection, but mixed GGO requires lobectomy, as the solid portion indicates invasiveness [3,5]. In other words, GGO is an important indicator of the extent of resection required and of the patient's prognosis. Moreover, better knowledge of this phenomenon could facilitate a deeper understanding of tumorigenesis in LUAD, as GGO adenocarcinoma is thought to be the precursor to invasive LUAD. However, because LUAD follows a stepwise progression, GGO-type LUAD cannot be analyzed in isolation [6].

LUAD is known to develop from atypical adenomatous hyperplasia, AIS, MIA, and invasive adenocarcinoma [6,7]. However, the causes of the development of AIS or MIA into an invasive form have not been determined. Many genetic factors, including p16 gene inactivation and mutations of epidermal growth factor receptor (*EGFR*), Kirsten rat sarcoma 2 viral oncogene homolog (*KRAS*), MNNG HOS transforming gene (*MET*), v-Raf murine sarcoma viral oncogene homolog B (*BRAF*), and tumor protein p53 (*TP53*), are known to be involved in the development of LUAD, but the detailed molecular events involved in each step have not been identified. Moreover, so-called driver mutations, including those of *EGFR*, *KRAS*, and *TP53*, have been identified frequently in the study of LUAD, but a significant number of LUAD patients have been found to lack these mutations [8,9].

Therefore, the purpose of this study was to investigate the genetic features of pure and mixed GGO LUAD in comparison to those of normal tissue through whole-exome sequencing (WES) of paired tissue (tumor and normal) DNA and to compare the data with non-GGO LUAD.

Methods

Ground-glass opacity and patients

Both pure GGO (which lacks a consolidation component) and mixed GGO (which contains both pure GGO and a consolidated region) were included in the study. Preoperative computed tomography images were reviewed by a single radiologist who was blinded to the pathological results and who classified the areas regarding whether they constituted GGO. All specimens were evaluated microscopically by a single pathologist who was blinded to the clinical and radiological data. Comprehensive histological subtyping was performed according to the eighth edition of the International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society classification system [10].

Patients with clinical T1N0 LUAD with GGO were included in the study. The control groups were patients of the same clinical stage (T1N0) who lacked GGO and exhibited consolidation only. Patients were excluded if they had clinical T2–4 cancer, N1–3 cancer, or any intrathoracic or extrathoracic metastases. Patients receiving neoadjuvant treatment and those with non-adenocarcinoma lung cancer were also excluded. The Institutional Review Board of Chungbuk National University Hospital approved this study (IRB approval no., 2017-04-018-004).

Whole-exome sequencing

Tumor and normal tissues were obtained by the surgeon in the operating room to guarantee the freshness of the tissue. The normal tissue samples were collected as far as possible from the tumor. Then, the tissues were confirmed by the pathologist to be tumor or normal cells. DNA was extracted using a Blood & Cell Culture DNA Mini Kit according the manufacturer's instructions (Qiagen, Hilden, Germany). It was quantitatively analyzed using a NanoDrop apparatus (Thermo Fisher Scientific, Waltham, MA, USA), and the integrity was assessed using 2% agarose electrophoresis with SYBR Gold (Thermo Fisher Scientific) staining. The genomic DNA of the tumors and matched normal tissues from all patients was captured using Agilent's in-solution enrichment methodology (SureSelect XT V5; Agilent, Santa Clara, CA, USA) with a biotinylated oligonucleotide probe library (SureSelect XT V5, 51 Mb; Agilent) followed by paired-end 101-base massively parallel sequencing on an Illumina HiSeq 4000 machine (Illumina, San Diego, CA, USA) [11].

Bioinformatics

The raw sequencing data (in the BCL file format) were converted to FASTQ files using Casava software (Illumina). Then, the Burrows-Wheeler alignment tool was utilized to map short reads to the human genome reference sequence (GRCh37). Duplicate read removal was performed using Picard (Broad Institute, Cambridge, MA, USA) and Samtools (Genome Research Limited, London, UK), and recalibration and variant calling were then performed using the Genome Analysis Toolkit (https://gatk.broadinstitute.org/hc/en-us). Variant annotations and effect predictions were subsequently performed with the SnpEff tool (https://pcingola.github.io/SnpEff/index.html).

The variations were identified by comparing the exome sequencing data of the tumor sample to the reference genome sequence derived from the paired normal tissue. Next, a Venn diagram summary of the unique variations was created to reveal any discrepancies in variations between GGO LUAD and the normal tissue and between non-GGO LUAD and the normal tissue in the paired samples [12]. The variations were filtered with regard to the consequence to the protein: start/stop gain/loss, splice, in-frame insertion/deletion, and frameshift variants. Using WES, the tumor mutational burden (TMB) was calculated based on the identified mutations in both tumor and germline DNA [13]. We focused on nonsynonymous muta-

https://doi.org/10.5090/kjtcs.20.013

tions among the tumor-specific functional mutations in The Cancer Genome Atlas cohort, since synonymous mutations could not be directly involved in creating neoantigens. Each mutation was classified into one of 3 domains: GGO-specific, non-GGO–specific, and common to both types. After the top 10 genes in each domain were selected, protein interaction network analysis was conducted with the Search Tool for the Retrieval of Interacting Genes/Proteins ver. 10 database (PMID: 25352553) using the Molecular Complex Detection plugin in Cytoscape 3.7.0 (The Cytoscape Consortium, San Diego, CA, USA), which is a visualization tool for integrating many molecular states, such as the expression level and interaction information, into a unified conceptual framework (PMID: 14597658). Finally, we examined the frequency of *EGFR*, *KRAS*, and *TP53* in each domain. Based on the results, we sequenced each mutation in the stepwise progression of LUAD to estimate its clinical significance.

Results

Patients and tumor mutational burden

In total, 20 samples were analyzed in this study. Of them, GGO (n=2 for the pure type and n=3 for the mixed type) and non-GGO LUAD were each present in 5 patients. These included 7 female patients, 4 of whom had GGO LUAD. No smokers were included in the current study, but 3 patients were ex-smokers (1 of whom had GGO LUAD). The mean size of the mass among the GGO cases was 10 mm, while that among the non-GGO cases was 15 mm. The GGO group consisted of 4 AIS and 1 MIA, while the non-GGO group consisted of 4 well-differentiated and 1 poorly-differentiated LUAD.

Tumors (n=10) and paired normal tissue (n=10) were used for analysis. In this study, 27,466 variations were identified in GGO and 27,376 in non-GGO. Of them, 227 functional variations were found in GGO, 212 in non-GGO, and 48 in common (Fig. 1). The Venn diagram summary showed that many discrepancies were present between the GGO and normal variations and between the non-GGO and normal variations in the paired samples. The TMB was 8.8 in the GGO group and 7.8 in the non-GGO group, with TMB defined as the number of mutations per megabase of DNA using the exome spans of coding sequences, which comprise less than 2% of the genome.

Gene alterations in each group

In the GGO domain, genes were selected as significant if they were found to have at least 2 mutations that were present in at least 2 samples. The top 10 most significant genes were *FCGBP*, *ANKRD36*, *C90RF50*, *CENPV*, *IGSF3*, *KIR2DS4*, *LGALS9B*, *PRR21*, *SFTPA1*, and *YBX2* (Fig. 2A). Among them, *FCGBP* was the most frequently mutated (with 4 mutations and 3 affected samples). Protein interaction network analysis was performed, but no clinically important relationships were found in the GGO domain.

Among the mutant genes commonly expressed in GGO and non-GGO LUAD, the top 10 most significant genes were *OR2T34*, *TCP10L2*, forkhead box Q1 (*FOXQ1*), *GBP4*, interferon regulatory factor 5 (*IRF5*), melanoma-associated antigen C1 (*MAGEC1*), *NUTM2F*, *PRR21*, *PRSS1*, and *PTPRT* (Fig. 2B). Among them, *OR2T34* was the most frequently-appearing gene (n=5, 50%), and *TCP10L2* (n=4, 40%) was the next most frequent. Protein interaction network analysis showed that *IRF5* and *GBP4* were related to the cellular response to cytokine stimuli (GO: 0071345) and the type I interferon signaling pathway (GO: 0060337) (Fig. 3A).

In the non-GGO domain, the top 10 most significant genes were cell division cycle 27 (*CDC27*), *NOTCH4*, *CES1*, *FAM58A*, *FRG2C*, *IGSF3*, *MIR4273*, *NUTM2G*, *OR2T27*, and *ZNF717* (Fig. 2C). Among them, *CDC27* and

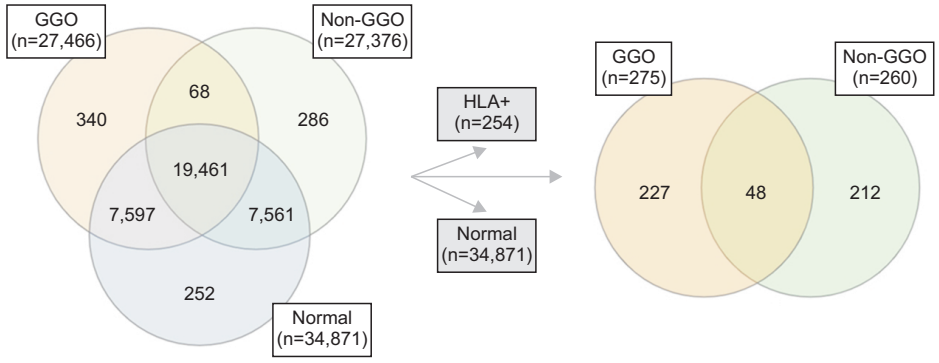
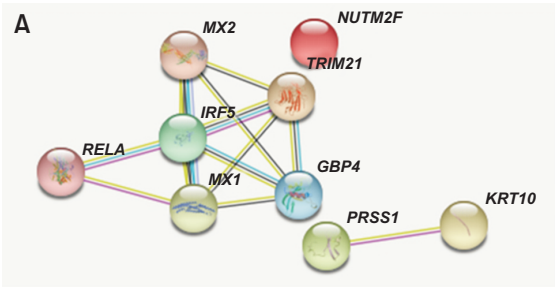


Fig. 1. Venn diagram of altered functional genes derived from whole-exome sequences. Many discrepancies were found among the GGO, non-GGO, and normal variations. The number of GGO-specific mutations was 227, compared to 212 in the non-GGO lung adenocarcinoma and 48 that were common to both tumor types, excluding normal variations and *HLA*-related genes. GGO, ground-glass opacity; HLA, human leukocyte antigen.

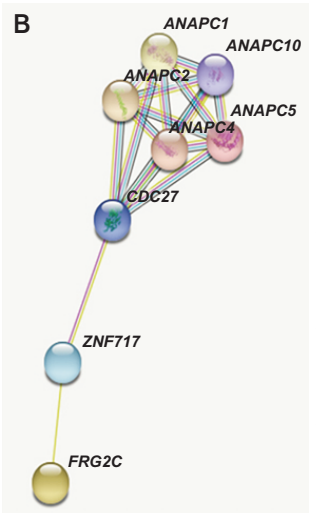
Fig. 2. Landscape of functional mutations in each domain. The top 10 mutations related to cellular functions were identified in each group. (A) *FCGBP* and *ANKRD36* were the most frequent gene alterations in the GGO group. (B) *OR2T34* and *TCP10L2* were the most frequent gene alterations common to both groups, but *FOXQ1*, *MAGEC1*, and *IFR5* have clinical implications in lung neoplasms. (C) *CDC27* and *NOTCH4* were the most frequent mutations in the non-GGO LUAD. GGO, ground-glass opacity; LUAD, lung adenocarcinoma; SNV, single nucleotide variant.

NOTCH4 were the most frequently identified genes (with 4 mutations and 2 affected samples). Based on protein interaction network analysis, *CDC27* and *ZNF717* were determined to be involved in anaphase-promoting complex-dependent catabolism (GO:0031145), cell division (GO:0051301), and the cell cycle (GO:0007049) (Fig. 3B).

https://doi.org/10.5090/kjtcs.20.013



Significant biological process	
GO:0071345	Cellular response to cytokine stimulus
GO:0060337	Type I interferon signaling pathway
GO:0051607	Defense response to virus
GO:0045087	Innate immune response
GO:0032495	Response to muramyl dipeptide



Significant biological process	
GO:0070979	Protein K11-linked ubiquitination
GO:0031145	Anaphase-promoting complex-dependent catabolic process
GO:0006508	Proteolysis
GO:0051301	Cell division
GO:0007049	Cell cycle

Fig. 3. Protein interaction network analysis. (A) *IRF5* (the most frequent mutation in common) and *GBP4* showed variable interactions with cytokine stimuli. (B) *CDC27* (the most frequent mutation in the non-GGO group) and *ZNF717* were determined to be associated with anaphase-promoting complex-dependent catabolism, cell division, and the cell cycle. GGO, ground-glass opacity.

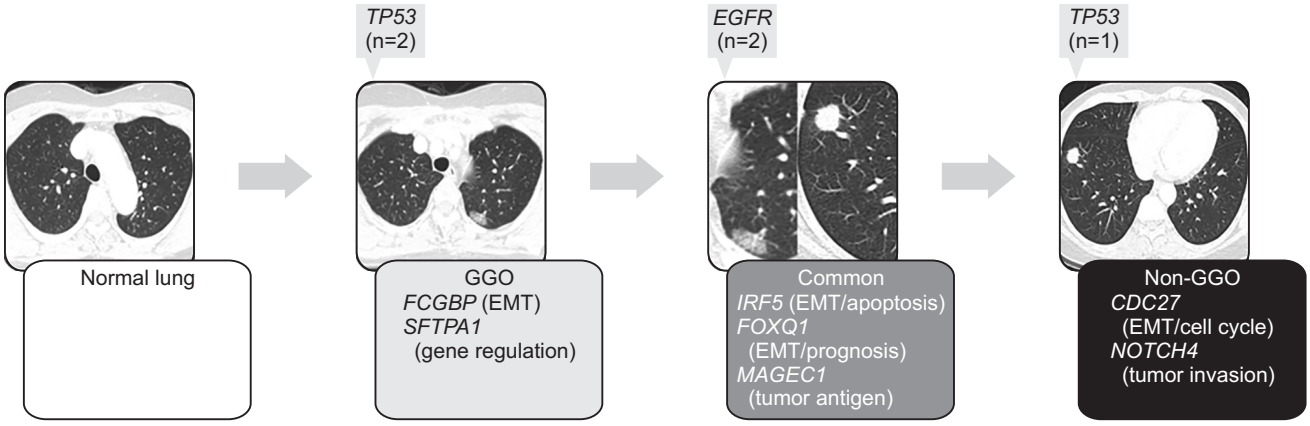


Fig. 4. Identified mutations and stepwise progression of lung adenocarcinoma. Mutations with clinical implications were identified according to the progression of lung adenocarcinoma. In the GGO step, mutations of *FCGBP*, *SFTPA1*, and *TP53* were identified, but these mutations lacked clinical implications. The identified *TP53* mutations were not identical between the GGO and non-GGO samples. *IRF5*, *FOXQ1*, *MAGEC1*, and *EGFR* were found to be mutated in both GGO and non-GGO samples; moreover, the mutations of *EGFR* were identical. In the non-GGO step, mutations of *CDC27*, *NOTCH4*, and *TP53* were detected. GGO, ground-glass opacity; EMT, epithelial-mesenchymal transition.

Although mutations of *EGFR*, *TP53*, and *KRAS* are clinically important, these genes were not present among the top 10 most common genes in the GGO or the non-GGO groups. Mutations of *EGFR* (n=2 in GGO and non-GGO; the mutation was identical) and *TP53* (n=2 in GGO and n=1 in non-GGO; all mutations were different) were observed, but no mutation of *KRAS* was found (Fig. 4).

Lung cancer progression according to genetic mutations

The identified mutations were arranged in the order of the stepwise progression of LUAD (Fig. 4). *FCGBP* and *SFTPA1* were positioned in the GGO step. The protein product of *FCGBP* (Fc fragment of immunoglobulin G binding protein) is known to be a key regulator of the transforming growth factor-1–induced epithelial-mesenchymal transition (EMT), and it has been found to be associated with

cancer progression, biological behaviors, and prognosis in gallbladder cancer [14]. The product of *SFTPA1* (surfactant, pulmonary-associated protein A1) plays a key role in innate lung host defense and gene regulation [15]. The expression levels of *SFTPA1* have been found to be much higher in the lung than in other tissues; moreover, this gene was observed to be strikingly downregulated in lung tumor tissues relative to the adjacent non-tumor tissues [16].

Among the commonly mutated genes, *IRF5*, *FOXQ1*, and *MAGEC1* were identified. *IRF5* encodes a member of the interferon regulatory factor family. This is a group of transcription factors with diverse roles, including in cell growth, differentiation, apoptosis, and immune system activity. *FOXQ1* is a member of the FOX gene family involved in embryonic development, cell cycle regulation, tissue-specific gene expression, cell signaling, and tumorigenesis. *MAGEC1* is a member of the melanoma antigen gene family. The proteins in this family are tumor-specific antigens that can be recognized by autologous cytolytic T lymphocytes.

Among the non-GGO–specific mutations, *CDC27* and *NOTCH4* were investigated. *CDC27* encodes a protein that is a component of the anaphase-promoting complex and has tetratricopeptide repeats, which are important for protein-protein interactions. This protein has been shown to interact with mitotic checkpoint proteins, and it may thus be involved in controlling the timing of mitosis. The Notch signaling pathway is an important form of cell-to-cell communication that plays a key role in regulating stem cell proliferation, differentiation, and apoptosis during embryonic development [17].

Discussion

This work produced 3 distinctive findings. First, the TMB was higher in the GGO than in the non-GGO tumor samples. A high TMB was found to be significantly associated with favorable outcomes in a previous study of 908 resected lung cancer tumor specimens, which were sequenced using a targeted panel of 1,538 genes [13]. It remains unclear why a high TMB is associated with better outcomes, but TMB has already been used in the patient selection process for a clinical trial of immune checkpoint inhibitors [18]. Further investigation of the TMB could elucidate the cause of the non-invasiveness of GGO. Second, driver mutations of *TP53* and *EGFR* were found in the current study, but they did not predominate, and no regularity was observed according to the increase in malignancy from GGO to non-GGO. As stated in many previous studies, it is clear that the above genes play an important role in lung cancer, but their role in the tumorigenesis of GGO or in pathologic T1 LUAD seems to be limited. Third, as the degree of malignancy increased, the mutation frequency of genes involved in the cell cycle and apoptosis also increased. In the GGO-specific stage, *FCGBP* and *SFTPA1* were identified, but they were not determined to be significant in lung cancer. However, among the mutations that were common to both tissue types or specific to the non-GGO stage, EMT- and tumor invasion-related genes were detected (Fig. 4).

All but the GGO-specific mutations had clinical significance. *FCGBP* and *SFTPA1* were identified in the GGO domain, but no meaningful research was available regarding lung cancer and protein interaction network analysis. *IRF5*, *FOXQ1*, and *MAGEC1*, which were common to both the GGO and non-GGO tumors, were clinically notable. *IRF5* was identified and validated as an EMT-regulating transcription factor in a study of the H358 non–small-cell lung cancer cell line [19]. In a study of *FOXQ1* and 4 common EMT indicator proteins (E-cad, MUC1, VIM, and S100A4), increases in *FOXQ1* expression were found to be correlated with changes in the expression of EMT indicators [20]. *MAGEC1*, which is normally expressed in male germ cells, is often expressed in tumors, including non–small-cell lung cancer; thus, it has been considered an attractive target for a cancer vaccine [21]. In the non-GGO samples, *CDC27* and *NOTCH4* were notable. *CDC27* is also involved in the EMT and has been found to promote factors that induce metastasis and invasion, but these findings have been made in the contexts of colorectal and stomach cancer [22,23]. Regarding *NOTCH4*, a study of its expression found that it was positively associated with tumor size, lymph node metastasis, and distant metastasis, and it was an independent prognostic factor for overall survival in lung cancer patients [24].

Mutations in the *EGFR* and *TP53* genes were found in the present study, but they were not predominant mutations in either the GGO or non-GGO groups. However, given that *EGFR* is expressed in both the GGO and non-GGO domains at the same time and that their mutations coincide completely, it is clear that the effect of *EGFR* on tumor progression is apparent in some GGO LUADs [25]. In addition, it should not be concluded that *TP53* is meaningless simply because its mutation was not found to be predominant in this study. Despite this lack of predominance and the fact that *TP53* does not fit into the stepwise progression schema in the study, it is reasonable to assume that this gene plays a specific role in both the GGO and

non-GGO domains [26].

This study is unique in its use of WES in the investigation of GGO and counterpart tumors. Due to cost limitations, most previous studies have been panel studies of limited number of genes. WES could provide additional data regarding unbiased pathogenic sequence mutations, including single-nucleotide variations and small insertion-deletions. Moreover, the landscape of mutations associated with LUAD exhibits racial and ethnic differences. Therefore, it is urgent to build up a database of GGO LUAD from different populations. Although this study was done with a limited sample size, its results could be a stepping stone to identifying the tumorigenesis of LUAD in Asian people.

In conclusion, GGO LUAD exhibited a higher TMB than non-GGO LUAD. No clinically meaningful mutations were found to be specific to GGO LUAD, but frequent mutations involving the EMT or the cell cycle were found in common and in the non-GGO domain. *EGFR* and *TP53* were found to have limited roles in this study, but further research is required. Overall, these findings could explain the non-invasiveness of GGO and may be useful in the search for targets for new medications.

Conflict of interest

No potential conflict of interest relevant to this article was reported.

Acknowledgments

The biospecimens used in the present study were provided by the Chungbuk National University Hospital, a member of the National Biobank of Korea, which is supported by the Ministry of Health, Welfare, and Family Affairs. All samples derived from the National Biobank of Korea were obtained with informed consent under institutional review board-approved protocols. The authors wish to thank Ms. Eun-Ju Shim from the National Biobank of Korea at Chungbuk National University Hospital for the sample preparations and her excellent technical assistance. The datasets generated and/or analyzed in the current study are available from the corresponding author upon reasonable request.

Funding

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Science, ICT & Future Planning (2017R1C1B5015969).

ORCID

Dohun Kim: https://orcid.org/0000-0001-8304-0232
Jong-Young Lee: https://orcid.org/0000-0002-0092-9958
Jin Young Yoo: https://orcid.org/0000-0003-0007-1960
Jun Yeun Cho: https://orcid.org/0000-0003-4270-413X

Authors' contributions

Dohun Kim designed the study and all experiments; Jong-Young Lee performed all experiments; Jin Young Yoo, Jun Yeun Cho and Dohun Kim collected patient samples; Jin Young Yoo and Jun Yeun Cho assisted with data collection; Jong-Young Lee analyzed the data; Jin Young Yoo, Jun Yeun Cho and Dohun Kim provided funding; Jong-Young Lee and Dohun Kim wrote the paper.

References

1. National Lung Screening Trial Research Team, Aberle DR, Adams AM, et al. *Reduced lung-cancer mortality with low-dose computed tomographic screening*. N Engl J Med 2011;365:395-409.
2. Aberle DR, DeMello S, Berg CD, et al. *Results of the two incidence screenings in the National Lung Screening Trial*. N Engl J Med 2013;369:920-31.
3. Kim D, Kim HK, Kim SH, et al. *Prognostic significance of histologic classification and tumor disappearance rate by computed tomography in lung cancer*. J Thorac Dis 2018;10:388-97.
4. Berry MF, Gao R, Kunder CA, et al. *Presence of even a small ground-glass component in lung adenocarcinoma predicts better survival*. Clin Lung Cancer 2018;19:e47-51.
5. Travis WD, Asamura H, Bankier AA, et al. *The IASLC Lung Cancer Staging Project: proposals for coding t categories for subsolid nodules and assessment of tumor size in part-solid tumors in the forthcoming eighth edition of the TNM classification of lung cancer*. J Thorac Oncol 2016;11:1204-23.
6. Noguchi M. *Stepwise progression of pulmonary adenocarcinoma: clinical and molecular implications*. Cancer Metastasis Rev 2010;29:15-21.
7. Travis WD, Brambilla E, Nicholson AG, et al. *The 2015 World Health Organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification*. J Thorac Oncol 2015;10:1243-60.
8. Jamal-Hanjani M, Wilson GA, McGranahan N, et al. *Tracking the evolution of non-small-cell lung cancer*. N Engl J Med 2017;376:2109-21.

9. Sakamoto H, Shimizu J, Horio Y, et al. *Disproportionate representation of KRAS gene mutation in atypical adenomatous hyperplasia, but even distribution of EGFR gene mutation from preinvasive to invasive adenocarcinomas*. J Pathol 2007;212:287-94.
10. Goldstraw P, Chansky K, Crowley J, et al. *The IASLC Lung Cancer Staging Project: proposals for revision of the TNM stage groupings in the forthcoming (eighth) edition of the TNM classification for lung cancer*. J Thorac Oncol 2016;11:39-51.
11. Gnirke A, Melnikov A, Maguire J, et al. *Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing*. Nat Biotechnol 2009;27:182-9.
12. Heberle H, Meirelles GV, da Silva FR, Telles GP, Minghim R. *InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams*. BMC Bioinformatics 2015;16:169.
13. Devarakonda S, Rotolo F, Tsao MS, et al. *Tumor mutation burden as a biomarker in resected non-small-cell lung cancer*. J Clin Oncol 2018;36:2995-3006.
14. Xiong L, Wen Y, Miao X, Yang Z. *NT5E and FcGBP as key regulators of TGF-1-induced epithelial-mesenchymal transition (EMT) are associated with tumor progression and survival of patients with gallbladder cancer*. Cell Tissue Res 2014;355:365-74.
15. Silveyra P, Floros J. *Genetic complexity of the human surfactant-associated proteins SP-A1 and SP-A2*. Gene 2013;531:126-32.
16. Peng L, Bian XW, Li DK, et al. *Large-scale RNA-Seq transcriptome analysis of 4043 cancers and 548 normal tissue controls across 12 TCGA cancer types*. Sci Rep 2015;5:13413.
17. Xiao W, Gao Z, Duan Y, Yuan W, Ke Y. *Notch signaling plays a crucial role in cancer stem-like cells maintaining stemness and mediating chemotaxis in renal cell carcinoma*. J Exp Clin Cancer Res 2017;36:41.
18. Hellmann MD, Ciuleanu TE, Pluzanski A, et al. *Nivolumab plus ipilimumab in lung cancer with a high tumor mutational burden*. N Engl J Med 2018;378:2093-104.
19. Du L, Yamamoto S, Burnette BL, et al. *Transcriptome profiling reveals novel gene expression signatures and regulating transcription factors of TGFβ-induced epithelial-to-mesenchymal transition*. Cancer Med 2016;5:1962-72.
20. Feng J, Zhang X, Zhu H, Wang X, Ni S, Huang J. *FoxQ1 overexpression influences poor prognosis in non-small cell lung cancer, associates with the phenomenon of EMT*. PLoS One 2012;7:e39937.
21. Sebastian M, Papachristofilou A, Weiss C, et al. *Phase Ib study evaluating a self-adjuvanted mRNA cancer vaccine (RNActive(R)) combined with local radiation as consolidation and maintenance treatment for patients with stage IV non-small cell lung cancer*. BMC Cancer 2014;14:748.
22. Qiu L, Tan X, Lin J, et al. *CDC27 induces metastasis and invasion in colorectal cancer via the promotion of epithelial-to-mesenchymal transition*. J Cancer 2017;8:2626-35.
23. Xin Y, Ning S, Zhang L, Cui M. *CDC27 facilitates gastric cancer cell proliferation, invasion and metastasis via twist-induced epithelial-mesenchymal transition*. Cell Physiol Biochem 2018;50:501-11.
24. Wang Y, Yang R, Wang X, et al. *Evaluation of the correlation of vasculogenic mimicry, Notch4, DLL4, and KAI1/CD82 in the prediction of metastasis and prognosis in non-small cell lung cancer*. Medicine (Baltimore) 2018;97:e13817.
25. Yang Y, Yang Y, Zhou X, et al. *EGFR L858R mutation is associated with lung adenocarcinoma patients with dominant ground-glass opacity*. Lung Cancer 2015;87:272-7.
26. Mao W, Wen Y, Lei H, et al. *Isolation and retrieval of extracellular vesicles for liquid biopsy of malignant ground-glass opacity*. Anal Chem 2019;91:13729-36.