



Characteristics of genomic mutations and signaling pathway alterations in thymic epithelial tumors

Weilin Yang^{1#}, Sai Chen^{2#}, Xinxin Cheng¹, Bo Xu¹, Huilan Zeng¹, Jianyong Zou³, Chunhua Su³, Zhenguang Chen^{1,3#^}

¹Department of Cardiothoracic Surgery of East Division, the First Affiliated Hospital, Sun Yat-sen University, Guangzhou, China; ²Center for Private Medical Service & Healthcare, the First Affiliated Hospital, Sun Yat-sen University, Guangzhou, China; ³Department of Thoracic Surgery, the First Affiliated Hospital, Sun Yat-sen University, Guangzhou, China

Contributions: (I) Conception and design: W Yang, Z Chen; (II) Administrative support: S Chen, Z Chen; (III) Provision of study materials or patients: W Yang, C Su, Z Chen; (IV) Collection and assembly of data: W Yang, S Chen, X Cheng, B Xu, H Zeng, J Zou, C Su; (V) Data analysis and interpretation: W Yang, X Cheng, Z Chen; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Dr. Zhenguang Chen, MD, PhD. Chief Surgeon, Department of Thoracic Surgery, and Department of Cardiothoracic Surgery of East Division, the First Affiliated Hospital, Sun Yat-sen University, No. 58, Zhongshan Road II, Guangzhou 510080, China. Email: chzheng@mail.sysu.edu.cn.

Background: To elucidate the mechanisms of thymic epithelial tumor (TET) canceration by characterizing genomic mutations and signaling pathway alterations.

Methods: Primary tumor and blood samples were collected from 21 patients diagnosed with TETs (thymoma and thymic cancer), 15 of whom were screened by nucleic acid extraction and whole exon sequencing. Bioinformatics was used to comprehensively analyze the sequencing data for these samples, including gene mutation information and the difference of tumor mutation burden (TMB) between thymoma and thymic carcinoma groups. We performed signaling pathway and functional enrichment analysis using the WebGestalt 2017 toolkit.

Results: *ZNF429* (36%) was the gene with the highest mutation frequency in thymic carcinoma. Mutations in *BAP1* (14%), *AB11* (7%), *BCL9L* (7%), and *CHEK2* (7%) were exclusively detected in thymic carcinoma, whereas *ZNF721* mutations (14%) and *PABPC1* (14%) were found exclusively in thymoma. The mean TMB values for thymic carcinoma and thymoma were 0.722 and 0.663 mutations per megabase (Mb), respectively, and these differences were not statistically significant. The ErbB signaling pathway was enriched in the thymoma and intersection groups, and pathways of central carbon metabolism in cancer, longevity regulating and MAPK signaling were only found in the thymoma group, while pathways in cancer (hsa05200) was found in the thymoma and thymic carcinoma groups.

Conclusions: Multiple differences in somatic genes and pathways have been identified. Our findings provide insights into differences between thymoma and thymic carcinoma that could aid in designing personalized clinical therapeutic strategies.

Keywords: Thymic epithelial tumors (TETs); thymoma; PD-L1 expression; ErbB signaling pathway; MAPK signaling pathways

Submitted Aug 31, 2021. Accepted for publication Nov 03, 2021.

doi: 10.21037/atm-21-5182

View this article at: <https://dx.doi.org/10.21037/atm-21-5182>

[^] ORCID: 0000-0002-8833-5686.

Introduction

As the most common anterior mediastinal tumors in adults, thymic epithelial tumors (TETs) account for approximately 0.2% to 1.5% of all malignancies (1,2). According to the World Health Organization (WHO) histologic classification of TETs, these tumors can be divided into types A, AB, B1, B2, B3, Thymic carcinoma, Squamous cell carcinoma, Undifferentiated carcinoma, Large-cell neuroendocrine carcinoma, NOS and Micronodular Thymoma (MNT) (3,4). Compared with thymoma, thymoma has the characteristics of inconspicuous lobular structure, larger nucleus, unbalanced nucleus to pulp ratio, and obvious heterogeneity (5). The most common tissue types of thymic carcinoma are squamous cell carcinoma and undifferentiated carcinoma (6).

Generally, thymomas are prone to local recurrence rather than metastasis. At present, the main treatment for thymoma is surgical resection with postoperative radiotherapy and chemotherapy (7,8). However, even with surgery, chemotherapy, and radiotherapy, thymic carcinoma is associated with a high risk of recurrence and death (9). As a result, the 5-year median survival rate of thymoma is better than that of thymic carcinoma, standing at 69% and 36%, respectively (10). Despite the mutation information and signaling pathways of thymic epithelial tumors have been studied (11), the differences in molecular characteristics and signaling pathway changes between these two tumor types are currently unknown.

Dozens of biomarkers related to checkpoint inhibitors have been identified to date. Among them, tumor mutation burden (TMB), programmed cell death-1 (PD-1) ligand 1 (PD-L1), and microsatellite instability/deficient mismatch repair (MSI/dMMR) have been verified by phase III clinical trials and are widely used in clinical practice (12-14). Although it has been reported that TMB-high (TMB-H) alone is not suitable for predicting the immunotherapeutic effect against solid tumors (15), TMB can serve as a biomarker for predicting PD-1/PD-L1 immune response (16,17). In adults, the TMB of thymic tumors is generally 0.48 mutations per megabase (18). PD-L1 is commonly expressed in TETs, including in 23–92% of thymomas and 36–80% of thymic carcinomas (19-25).

Multiple studies have investigated the relationship between PD-L1 expression and survival/prognosis in patients with TETs; however, their results have been conflicting. In their study, Padda *et al.* reported patients with high PD-L1 expression to have poor overall survival (19). However, Weissferdt *et al.* failed to find a correlation between PD-L1 expression and prognosis (25), while

Arbour *et al.* and Yokoyama *et al.* reported a good prognosis in patients with high PD-L1 expression (22,26). Studies have also shown complex signaling pathway crossover in the tumor canceration process (27), although few studies have investigated signaling pathways in the process of thymoma canceration.

To better understand the potential mechanisms and biological characteristics of TETs, next-generation sequencing (NGS) and signal pathway analyses of TET samples were conducted to investigate somatic gene mutations, TMB, the transition/transversion (TI/TV) ratio, and mutated gene signaling pathways in patients with thymoma and thymic carcinoma. Our findings may provide effective guidance for clinical intervention, thereby improving the long-term prognosis of patients with TETs. We present the following article in accordance with the MDAR checklist (available at <https://dx.doi.org/10.21037/atm-21-5182>).

Methods

Patient specimen acquisition

Twenty-one patients with TETs including 9 with thymomas and 12 with thymic carcinomas were enrolled into this study. PD-L1 expression was detected in all 21 patients, and whole-exome sequencing (WES) was performed for 15 patients.

Slide-mounted tumor specimens were classified by hematoxylin and eosin (H&E) staining according to the 2010 WHO classification of digestive system tumors. The included tumor tissue samples were histologically confirmed as adenocarcinoma by two molecular pathologists and shown to consist of more than 70% tumor cells.

This study was conducted in accordance with the Helsinki Declaration (as revised in 2013) (28), and the protocol was approved by the Ethics Committee of the First Affiliated Hospital of Sun Yat-sen University (No. [2017]174). Informed written consent was obtained from each patient before their inclusion in the study.

DNA extraction and qualification

Tumor tissue sections were subjected to DNA extraction using a DNeasy Blood and Tissue Kit (69504; Qiagen, Venlo, Netherlands) in line with the manufacturer's instructions. The extracted DNA was characterized using an Agilent Bioanalyzer (Santa Clara, CA, USA). Germline mutations were determined by reference to matched leukocyte DNA.

Target gene sequencing and WES

A NadPrep DNA Library Preparation Kit (#1002101; Nanodigmbio, Nanjing, China) was used to prepare libraries for the Illumina sequencing platform, which is equipped with a unique dual index (UDI) adapter that can mitigate label-skipping and label-mismatch issues. A NadPrep Hybrid Capture Reagents kit (#1005101; Nanodigmbio) was used for hybridization capture-based target enrichment.

Variant annotation analysis and visual mapping

False-positive results were excluded through realigning split-reads to hg19/GRCh37 using Blast. The computational efficiency of WES data capture was assessed using the Picard tool (<http://broadinstitute.github.io/picard/command-line-overview.html#CollectHs-Metrics>), which filters low-quality bases, repetitive bases, bases that deviate from the target, and ends that overlap with reads at both ends due to short inserts. Using this tool, the strictest depth distribution was calculated. For the purpose of quality control, sequencing data were considered to be qualified if the mean bait coverage exceeded 100. To improve the alignment of individual reads, further local rearrangements were performed using SpeedSeq (29). Somatic mutations were identified using MuTect (30). ANNOVAR and Oncotator were used to annotate genetic variant data, which were converted to a MAF (Mutation Annotation Format) file using the maf tool (31). Cancer driver genes were analyzed using the Integrative OncoGenomics (IntOgen) platform, running Oncodrive FM and Oncodrive CLUST (32). The landscape of the top driver mutation spectrum for tumors predicted by IntOgen, including mutation rate and mutation subclass/subtypes (Oncodrive FM $P \leq 0.1$), was visualized with R Script.

Pathway enrichment analysis

Differences in the frequency of somatic gene mutations between the thymoma and thymic carcinoma groups were illustrated using Wayne diagrams, as some gene mutations occurred only in one group, while others occurred in both groups. WebGestalt (WEB-based Gene SeT AnaLysis Toolkit, <http://www.webgestalt.org>) was used to analyze the canonical pathways associated with genes containing single-nucleotide variations (SNVs). P values were calculated based on a hypergeometric distribution, and the Benjamini-Hochberg method was adopted for false

discovery rate correction (33).

Statistical analysis

Correlations of clinical and biological variables were analyzed using the Statistical Package for the Social Sciences (SPSS) 22.0 (IBM Corp., Armonk, NY, USA). Where necessary, non-parametric tests, such as Welch's *t*-test and Wilcoxon signed-rank test, in the R package ggpubr were used (34).

Results

Patient characteristics

To better understand the molecular mechanisms underlying the genesis of TETs, fresh frozen tumor tissues and corresponding blood samples were selected from 15 of the 21 patients with TETs in the study and were subjected to bioinformatics analyses according to strict standards. Clinicopathological assessment indicated that there were 9 cases (S01, S04, S05, S06, S10, S12, S13, S14, and S15) of thymic carcinoma, of which one case (S06) also had a lung metastasis. The other 12 cases (S02, S03, S07, S08, S09, S11, S16, S17, S18, S19, S20, and S21) were thymomas. Cases S16, S17, S18, S19, S20, and S21 had a tumor cell content of less than 50%; WES was not performed on these samples. Detailed clinical characteristics of the patients are shown in *Table 1*. The PD-L1 positive rate was 33.3% (3/9) and 58.3% (7/12) in the thymic carcinoma and thymoma groups, respectively (*Table 1* and *Figure 1*). The detailed clinical characteristics of the patients are shown in *Table S1*.

Identifying somatic mutations in 15 patients with TETs

NGS sequencing of DNA from 15 primary tumor tissues and matched blood samples was performed, and several somatic mutations were annotated using MuTect and Somatic Indel Detector. The average depth of WES of the tumor samples was 346.06x, and the average depth of sequencing of the control samples was 127.96x. A total of 1,779 nonsynonymous SNVs were detected, including 912 in thymic carcinomas and 867 in thymomas (<https://cdn.amegroups.com/static/public/ATM-21-5182-1.xlsx>).

As shown in *Figure 2*, case S08 had the most SNVs, followed by case S03. The top 30 driver genes were ranked based on the frequency of their somatic mutation, which showed that *ZNF429* (36%) had the highest mutation frequency. The most common mutation type was missense

Table 1 Patient characteristics

Characteristic	No. of cases
Total number	n=21
Age, years, mean [range]	51 [21–69]
Sex, n (%)	
Male	15 (71.4)
Female	6 (28.6)
Pathological type, n (%)	
Thymic carcinoma	9 (42.9)
Thymoma	12 (57.1)
PD-L1 positivity, n (%)	
Thymic carcinoma	3 (33.3)
Thymoma	7 (58.3)

mutations, followed by nonsense, splice site, and other mutations. Among the top 30 driver genes, *BAP1* (14%), *ABII* (7%), *BCL9L* (7%), and *CHEK2* (7%) mutations were detected exclusively in thymic carcinoma, while *ZNF721* (14%) and *PABPC1* (14%) mutations were found exclusively in thymoma. *ZNF429*, *ZNF208*, *BAP1*, *ERBB4*, *GNAQ*, *ABII*, *BCL9L* were common somatic mutations in thymic carcinoma subgroup compared with healthy people.

The TMB was calculated using only somatic nonsynonymous mutations. The mean TMB values were 0.722 and 0.663 mutations per Mb in the thymic carcinoma and thymoma groups, respectively, and the difference between the groups was not statistically significant ($P=0.817$, Welch's *t*-test; *Figure 3A*). Notably, the TMB value in this study (0.698) was significantly higher than that reported in The Cancer Genome Atlas thymoma (TCGA-THYM) (0.488) ($P=0.0038$, Wilcoxon signed-rank test; *Figure 3B*).

In principle, all mutation types (e.g., substitutions, indels, and rearrangements) and any accessory mutation characteristic (e.g., the sequence context of the mutation or the transcriptional chain where the mutation occurred) can be incorporated into the set of features by which a mutational signature is defined. Mutational signatures were extracted using base substitutions, and six classes of substitutions were assigned according to the pyrimidine of the mutated Watson-Crick base pair: C > A, C > G, C > T, T > C, T > A, and T > G. The six mutation types were then compared with the TCGA-THYM database, and the proportions of these six mutation replacement types were

found to be roughly the same. Among all the substitutions, C > T had the highest mutation percentage, which differed significantly from that for the TCGA-THYM database ($P=0.0014$). The T > G, C > A, and T > A substitutions in this study also differed significantly from those in the TCGA-THYM study (*Figure 3C,3D*).

Comparison of signaling pathway alterations between thymoma and thymic carcinoma

For further characterization the functions of mutational genes and pathways involved in TET, the PANTHER classification system, an ontology-based pathway database, was coupled with data analysis tools (35). Using pathological diagnosis results, a Venn diagram was created, which divided the driver mutant genes detected by WES into three clusters: thymic carcinoma (38 genes), thymoma (38 genes), and the intersection of the two (15 genes) (*Figure 4A*).

In thymic carcinoma, mutant somatic genes were mainly enriched in chronic myeloid leukemia (hsa05220), the ErbB signaling pathway (hsa04012), and the T-cell receptor signaling pathway (hsa04660) (*Figure 4B*). Meanwhile, altered somatic genes in thymoma were mainly enriched in central carbon metabolism in cancer (hsa05230), acute myeloid leukemia (hsa05221), MAPK signaling pathways (hsa04010), and longevity-regulating pathways (hsa04213) (*Figure 4C*). The shared driver gene enrichment pathways at the intersection of the thymic carcinoma and thymoma clusters mainly included lysine degradation (hsa00310), endocrine and other factor-regulated calcium reabsorption (hsa04961), long-term depression (hsa04730), long-term potentiation (hsa04720), gastric acid secretion (hsa04971), and the ErbB signaling pathway (hsa04012) (*Figure 4D*).

Discussion

TET, a rare malignant tumor of the chest, is the most common anterior mediastinal neoplasm in adults, with an incidence rate of 1.5–3.2 cases per million people per year (36). Among TETs, thymomas are often uniquely associated with autoimmune diseases, whereas thymic carcinomas exhibit more invasive characteristics in clinic. A study by Radovich *et al.* reported a strong association of broad histological subtypes (A, AB, B, and thymic carcinoma) with multiple classes of aberrations occurring at different levels (18). The etiology of TETs is unknown, largely due to the limited knowledge of the genomic underpinnings of thymoma and thymic carcinoma.

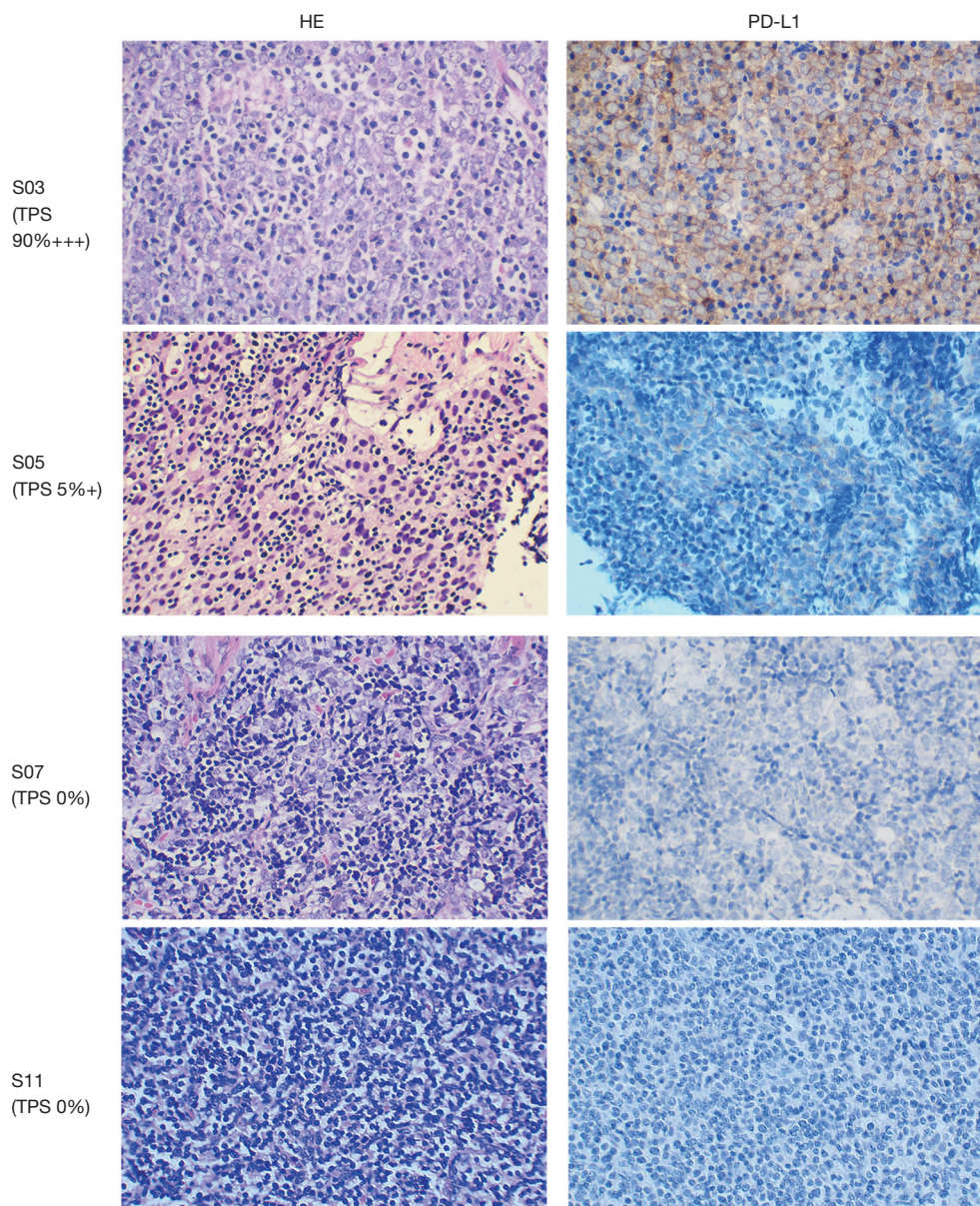


Figure 1 Pictures of hematoxylin and eosin staining and programmed death-ligand 1 (PD-L1) detection of thymic epithelial tumor samples (all images were $\times 400$ magnification fields).

In this study, WES of 15 tumor-normal paired samples revealed some high-frequency mutation driver genes that have been reported in previous studies on TETs, including *BAP1*, *ERBB4*, *HRAS*, and *KIT* (37,38). *BAP1* and *BCL9L*, as oncogenes, promote the growth, invasion and lung metastasis of breast cancer by interacting with β -catenin and Pygopus (39-41). *ABI1* inhibits epithelial-mesenchymal transformation by inhibiting atypical WNT signal activation, providing a new mechanism for the

inhibition of prostate tumors (42). Analysis further showed that the *ZNF429* (36%) and *ZNF208* (29%) genes had the highest frequency of mutations (Figure 2). The *ZNF429* gene was reported to be significantly associated with acute postoperative pain in patients with breast cancer, while *ZNF208* polymorphisms were found to be associated with ischemic stroke in a southern Chinese Han population (43,44). Research has found that PABPC1 [poly(A) binding protein cytoplasmic 1] is directly involved in the



Figure 2 Somatic mutations in thymic epithelial tumor (TET) samples. Top: variations in the numbers of mutations in each sample; middle: mutated genes and mutation types; right: frequency of mutations in somatic genes in TET samples.

tumorigenesis of gastric cancer through its promotion of cell proliferation; consistent with this, downregulated *PABPC1* expression is associated with tumor progression and a poor prognosis in colorectal cancer (45,46). Our analysis of the TCGA-THYM dataset further revealed that the genes with the highest mutation frequency were *GTF2I* (50%), *HRAS* (8%), and *MUC16* (7%) (Figure 5). Radovich *et al.* showed that mutated *GTF2I* is a thymoma-specific oncogene, and that cells expressing mutant *GTF2I* also highly express genes involved in the Wnt and Sonic hedgehog (SHH) signaling pathways (18). Perplexingly, the current study did not detect mutations in the *GTF2I* gene in thymoma or thymic carcinoma samples, which is possibly due to insufficient sample size. *MUC16*, a mucin expressed at high levels on the surface of epithelial ovarian

tumor cells, exerts similar effects to *MUC1*, which binds or aggregates neutrophils and other cell types to provide antitumor immune protection (47,48).

PD-L1 expression has been used as a tumor cell marker in a number of clinical trials and has been approved by FDA (the food and drug administration) (49). Positive expression of PD-L1 in tumors is considered to indicate patients who are more likely to respond to immune checkpoint inhibitors treatment (50). However, some patients who test positive for PD-L1 may fail to respond to treatment, and—more importantly—some patients who test negative may respond, making this an imperfect biomarker (51). Multiple studies have reported conflicting data on the relationship between PD-L1 expression and survival/prognosis in patients with TETs (19,22,25,26). In the current study, the rates of PD-L1

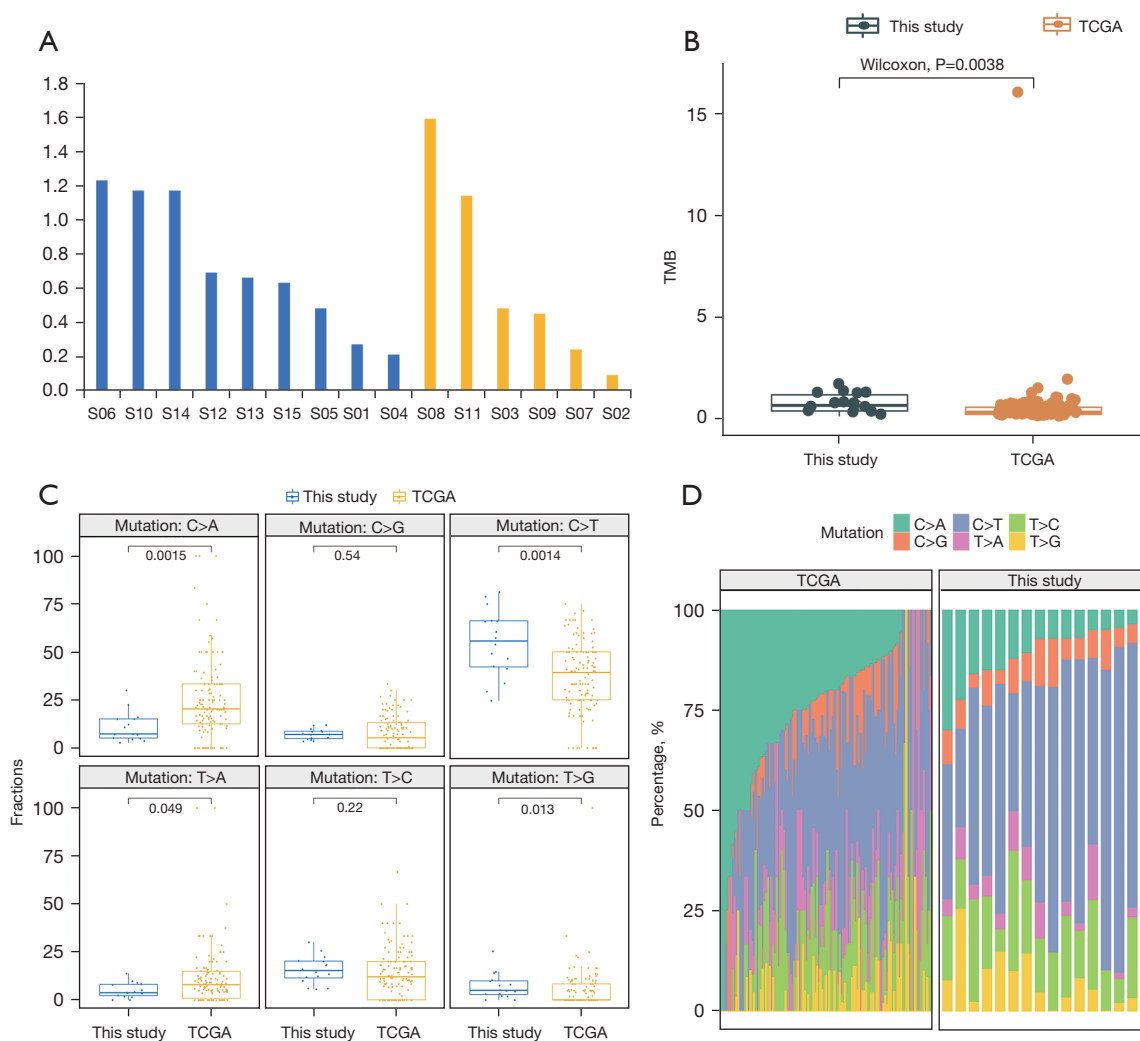


Figure 3 Comparative analysis of differences in tumor mutational burden (TMB). (A) TMB in thymic carcinoma and thymoma samples. (B) Comparison of TMB between the data in this study and The Cancer Genome Atlas thymoma (TCGA-THYM) database. (C) Comparative analysis of the relative transition/transversion (TI/TV) ratio between this study and the TCGA-THYM study. (D) Transition and transversion proportions of six nucleotide changes. The stacked proportion bar chart is sorted according to the increase in TI/TV fraction.

positivity in thymoma and thymic carcinoma (33.3% and 58.3%, respectively) were consistent with the positivity-rate range reported in the literature (19-25).

High TMB is a characteristic associated with responsiveness to immunotherapy (52). However, a high TMB value (≥ 20 mutations per Mb) is not suitable for predicting the immunotherapeutic effect in all solid tumors (14). Establishing the high TMB threshold for different cancer types may require additional clinical studies and information statistics of a larger number of patients. The TMB value in this study was significantly higher than that in the TCGA-THYM dataset, which suggests that the

patients in our study would be more likely to benefit from immunotherapy (Figure 3).

Because of the differences between our findings regarding TMB and those reported by the TCGA-THYM dataset, we also analyzed the types of mutation, as defined by the TI/TV ratio. Significant differences existed between our study and the TCGA data with respect to C > A, C > T, T > A, and T > G mutations; however, no significant difference was found for C > G and T > C mutations (Figure 3). Mutational signatures can be understood as different mutation processes that often generate different combinations of mutation types (53). Thousands of somatic

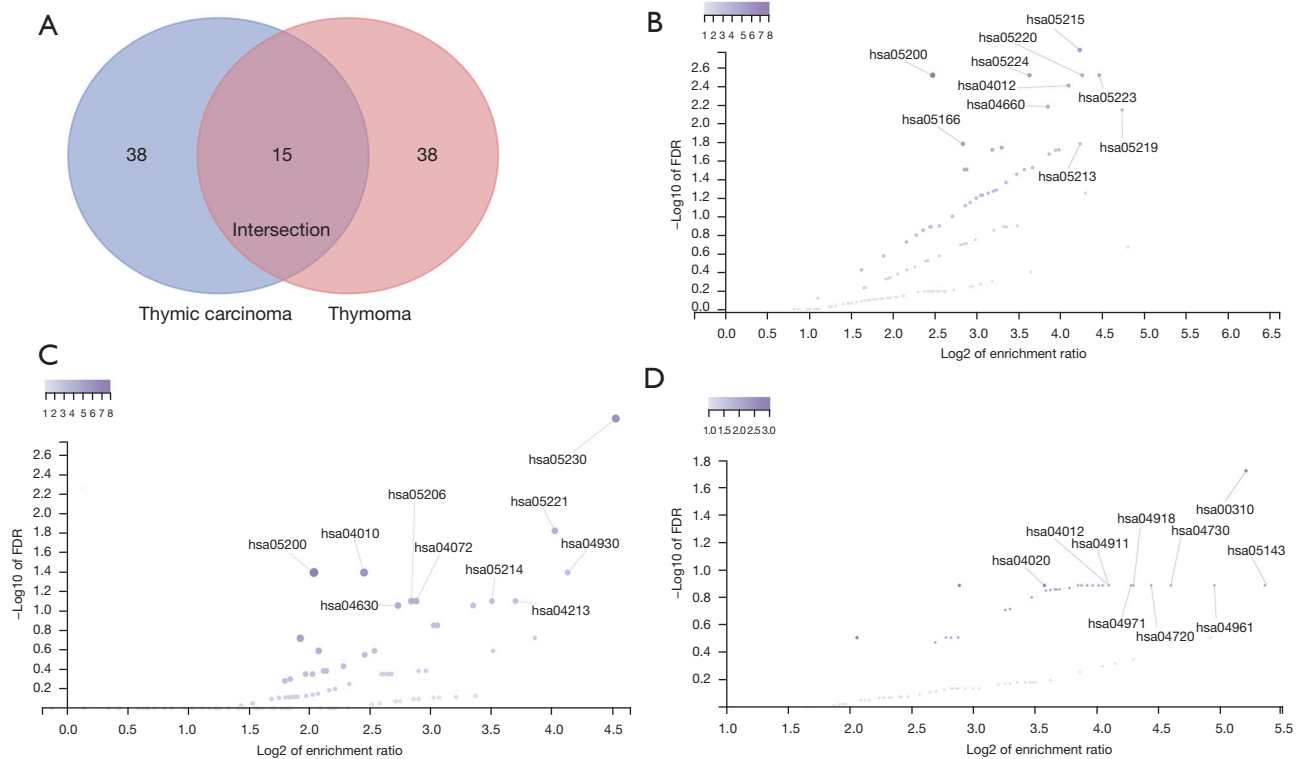


Figure 4 Analysis of pathways involving nonsynonymous genes in thymic epithelial tumors. (A) Venn diagram showing gene distribution in three clusters. (B) Volcano plot of uniquely mutated driver gene enrichment pathways in thymic carcinoma. (C) Volcano plot of uniquely mutated driver gene enrichment pathways in thymoma. (D) Volcano plot of shared driver gene enrichment pathways in the intersection group.

mutations can be identified in a single cancer sample, thus making it possible to decipher the mutant signature, even if several mutations are operative (54). The C > A mutational signature has been shown to be associated with tobacco smoking, whereas C > T mutations are induced by ultraviolet light (55).

Single or multiple genes alternation and expression level change usually response to pathway change and complex cross-talk between different signaling pathways simultaneously, such as the angiogenesis and notch signaling pathways (56). Cancers are not driven by single gene mutations or expression changes but by coordinated changes affecting multiple signaling pathways (57,58). ErbB (erythroblastic oncogene B), which is also called the epidermal growth factor receptor (EGFR), is generally expressed on the surface of epithelial cells, but is often overexpressed in certain tumor cells (59). Overexpression of EGFR is associated with tumor cell metastasis and invasion, and a poor prognosis (60). In the current study, we found enrichment of the ErbB and T-cell signaling pathways

in a high proportion of samples in the thymic carcinoma group and that were also found at the intersection of gene clusters (Figure 4B,4C). In the thymoma group, prominent pathways included longevity-regulating pathways (hsa04213) and central carbon metabolism in cancer (hsa05230) (Figure 4C). Cell immortalization is an important stage in tumorigenesis (61). In mammals, calorie restriction (CR) is considered to be the most reliable and repeatable longevity-promoting strategy. Among the pathways involved in the regulation of the CR effect are the insulin-like growth factor (IGF-1)/insulin signaling pathway, the sirtuin pathway, the adenosine monophosphate (AMP)-activated protein kinase (AMPK) pathway, and the target of rapamycin (TOR) pathway (62). These pathways are believed to respond to CR via a common mechanism by which autophagy, stress defense mechanisms, and survival pathways are activated, and proinflammatory mediators and cell growth become weaker, thereby promoting cell health and, ultimately, longevity (63). Anti-PD-1/PD-L1 and other immunotherapeutic drugs have good efficacy and safety in

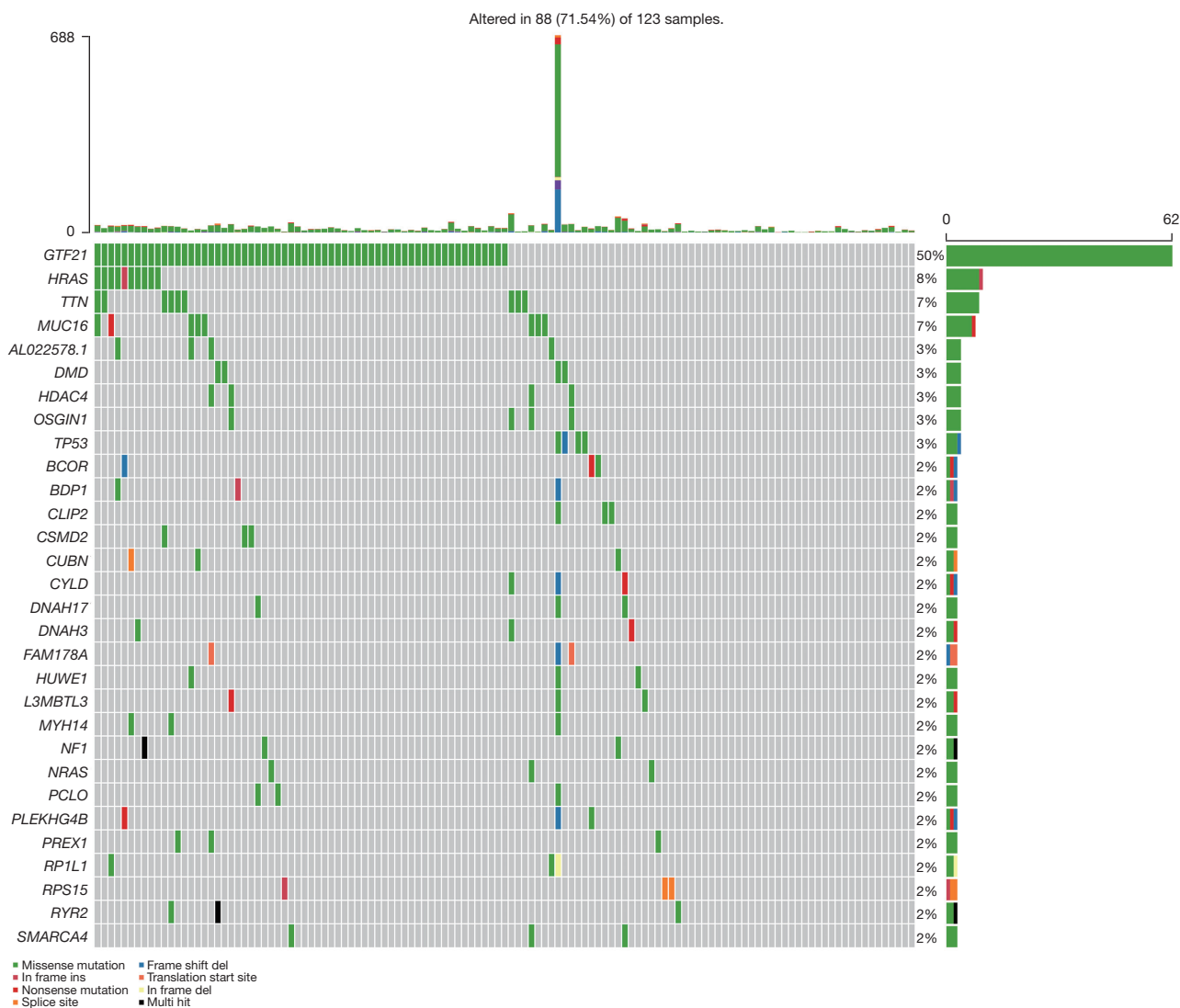


Figure 5 Mutational landscape of somatic alterations in The Cancer Genome Atlas thymoma (TCGA-THYM) database.

the treatment of solid tumors such as melanoma, non-small cell lung cancer, and head and neck cancer (64). At present, some domestic studies suggested that the expression of PD-L1 was related to the pathological classification of thymus, especially in type B2/B3 thymoma and thymic carcinoma, which indicated that the expression of PD-L1 was related to the prognosis of thymic cancer (65). However, whether patients with thymic cancer can be determined to benefit from immunotherapy requires more clinical research. In short, the carcinogenesis of thymic epithelial cells may involve benign or malignant mutations of multiple genes or cross feedback of signaling pathways. In the treatment of thymic epithelial tumors, we must not only use

surgical resection or adjuvant radiotherapy and chemotherapy based on the pathological staging of the patient, but also pay attention to the comprehensive treatment based on the molecular classification of the patient.

Conclusions

The present study has analyzed somatic gene mutations, TMB, and TI/TV ratios in thymoma and thymic carcinoma using bioinformatics, and has also expounded the differences between our findings and those of the TCGA-THYM dataset. Drive genes of *ZNF429*, *BAP1*, *ABI1*, *BCL9L*, *CHEK2*, *ZNF721* and *PABPC1* are expressed to varying

degrees in thymoma or thymic carcinoma, respectively. Enrichment of the longevity-regulating, MAPK, and ErbB signaling pathways in thymoma still remains to be further investigated. An in-depth understanding of the differences in somatic genes and pathways between thymoma and thymic cancer will help to design personalized clinical therapy for TET patients.

Acknowledgments

Shanghai Tongshu Biotechnology Co., Ltd. provided support for this study.

Funding: None.

Footnote

Reporting Checklist: The authors have completed the MDAR checklist. Available at <https://dx.doi.org/10.21037/atm-21-5182>

Data Sharing Statement: Available at <https://dx.doi.org/10.21037/atm-21-5182>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://dx.doi.org/10.21037/atm-21-5182>). The authors report that the Shanghai Tongshu Biotechnology Co., Ltd. provided support for this study (The unit performed experimental program operation and technical support). The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The protocol was approved by the Ethics Committee of the First Affiliated Hospital of Sun Yat-sen University (No. [2017]174). Informed written consent was obtained from each patient before their inclusion in the study.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the

original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- de Jong WK, Blaauwgeers JL, Schaapveld M, et al. Thymic epithelial tumours: a population-based study of the incidence, diagnostic procedures and therapy. *Eur J Cancer* 2008;44:123-30.
- Engels EA. Epidemiology of thymoma and associated malignancies. *J Thorac Oncol* 2010;5:S260-5.
- Imbimbo M, Ottaviano M, Vitali M, et al. Best practices for the management of thymic epithelial tumors: A position paper by the Italian collaborative group for ThYmic MalignanciEs (TYME). *Cancer Treat Rev* 2018;71:76-87.
- Travis WD, Brambilla E, Burke AP, et al. Introduction to The 2015 World Health Organization Classification of Tumors of the Lung, Pleura, Thymus, and Heart. *J Thorac Oncol* 2015;10:1240-2.
- Sharma S, Dawson L. A Rare Tumor with a Very Rare Initial Presentation: Thymic Carcinoma as Bone Marrow Metastasis. *Case Rep Pathol* 2017;2017:6497376.
- Yang X, Zhuo M, Shi A, et al. Optimal first-line treatment for advanced thymic carcinoma. *Thorac Cancer* 2019;10:2081-7.
- Girard N, Mornex F, Van Houtte P, et al. Thymoma: a focus on current therapeutic management. *J Thorac Oncol* 2009;4:119-26.
- Li J, Liu Q, Zheng Z, Li S. Prolonged survival time with surgical therapy in different types of thymoma: an analysis based on Surveillance Epidemiology and End Results Database. *Transl Cancer Res* 2020;9:5893-9.
- Eng TY, Fuller CD, Jagirdar J, et al. Thymic carcinoma: state of the art review. *Int J Radiat Oncol Biol Phys* 2004;59:654-64.
- Scorsetti M, Leo F, Trama A, et al. Thymoma and thymic carcinomas. *Crit Rev Oncol Hematol* 2016;99:332-50.
- Chen K, Che J, Zhang X, et al. Next-generation sequencing in thymic epithelial tumors uncovered novel genomic aberration sites and strong correlation between TMB and MSH6 single nucleotide variations. *Cancer Lett* 2020;476:75-86.
- Pietrantonio F, Loupakis F, Randon G, et al. Efficacy and Safety of Immune Checkpoint Inhibitors in Patients with Microsatellite Instability-High End-Stage Cancers and Poor Performance Status Related to High Disease Burden.

- Oncologist 2020;25:803-9.
13. Yang Q, Xu Z, Zheng L, et al. Multimodal detection of PD-L1: reasonable biomarkers for immune checkpoint inhibitor. *Am J Cancer Res* 2018;8:1689-96.
 14. Feng L, Hong S, Gao J, et al. Whole-Exome Sequencing Characterized the Landscape of Somatic Mutations and Pathways in Colorectal Cancer Liver Metastasis. *J Oncol* 2019;2019:2684075.
 15. Xu Z, Dai J, Wang D, et al. Assessment of tumor mutation burden calculation from gene panel sequencing data. *Onco Targets Ther* 2019;12:3401-9.
 16. Qin BD, Jiao XD, Zang YS. Tumor mutation burden to tumor burden ratio and prediction of clinical benefit of anti-PD-1/PD-L1 immunotherapy. *Med Hypotheses* 2018;116:111-3.
 17. Lu S, Stein JE, Rimm DL, et al. Comparison of Biomarker Modalities for Predicting Response to PD-1/PD-L1 Checkpoint Blockade: A Systematic Review and Meta-analysis. *JAMA Oncol* 2019;5:1195-204.
 18. Radovich M, Pickering CR, Felau I, et al. The Integrated Genomic Landscape of Thymic Epithelial Tumors. *Cancer Cell* 2018;33:244-258.e10.
 19. Padda SK, Riess JW, Schwartz EJ, et al. Diffuse high intensity PD-L1 staining in thymic epithelial tumors. *J Thorac Oncol* 2015;10:500-8.
 20. Katsuya Y, Fujita Y, Horinouchi H, et al. Immunohistochemical status of PD-L1 in thymoma and thymic carcinoma. *Lung Cancer* 2015;88:154-9.
 21. Yokoyama S, Miyoshi H, Nishi T, et al. Clinicopathologic and Prognostic Implications of Programmed Death Ligand 1 Expression in Thymoma. *Ann Thorac Surg* 2016;101:1361-9.
 22. Yokoyama S, Miyoshi H, Nakashima K, et al. Prognostic Value of Programmed Death Ligand 1 and Programmed Death 1 Expression in Thymic Carcinoma. *Clin Cancer Res* 2016;22:4727-34.
 23. Marchevsky AM, Walts AE. PD-L1, PD-1, CD4, and CD8 expression in neoplastic and nonneoplastic thymus. *Hum Pathol* 2017;60:16-23.
 24. Higuchi R, Goto T, Hirotsu Y, et al. PD-L1 Expression and Tumor-Infiltrating Lymphocytes in Thymic Epithelial Neoplasms. *J Clin Med* 2019;8:1833.
 25. Weissferdt A, Fujimoto J, Kalhor N, et al. Expression of PD-1 and PD-L1 in thymic epithelial neoplasms. *Mod Pathol* 2017;30:826-33.
 26. Arbour KC, Naidoo J, Steele KE, et al. Expression of PD-L1 and other immunotherapeutic targets in thymic epithelial tumors. *PLoS One* 2017;12:e0182665.
 27. Sanchez-Vega F, Mina M, Armenia J, et al. Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell* 2018;173:321-337.e10.
 28. World Medical Association Inc. Declaration of Helsinki. Ethical principles for medical research involving human subjects. *J Indian Med Assoc* 2009;107:403-5.
 29. Chiang C, Leyer RM, Faust GG, et al. SpeedSeq: ultra-fast personal genome analysis and interpretation. *Nat Methods* 2015;12:966-8.
 30. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 2011;43:491-8.
 31. Mayakonda A, Lin DC, Assenov Y, et al. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res* 2018;28:1747-56.
 32. Gonzalez-Perez A, Perez-Llamas C, Deu-Pons J, et al. IntOGen-mutations identifies cancer drivers across tumor types. *Nat Methods* 2013;10:1081-2.
 33. Li A, Barber RF. Multiple testing with the structure-adaptive Benjamini-Hochberg algorithm. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2019;81:45-74.
 34. Kassambara A. ggpubr: 'ggplot2' Based Publication Ready Plots 2018. Available online: <http://www.sthda.com/english/rpkgs/ggpubr>
 35. Mi H, Muruganujan A, Casagrande JT, et al. Large-scale gene function analysis with the PANTHER classification system. *Nat Protoc* 2013;8:1551-66.
 36. Janik S, Bekos C, Hacker P, et al. Follistatin impacts Tumor Angiogenesis and Outcome in Thymic Epithelial Tumors. *Sci Rep* 2019;9:17359.
 37. Arenzana TL, Lianoglou S, Seki A, et al. Tumor suppressor BAP1 is essential for thymic development and proliferative responses of T lymphocytes. *Sci Immunol* 2018;3:eaal1953.
 38. Enkner F, Pichlhöfer B, Zaharie AT, et al. Molecular Profiling of Thymoma and Thymic Carcinoma: Genetic Differences and Potential Novel Therapeutic Targets. *Pathol Oncol Res* 2017;23:551-64.
 39. Qin J, Zhou Z, Chen W, et al. BAP1 promotes breast cancer cell proliferation and metastasis by deubiquitinating KLF5. *Nat Commun* 2015;6:8471.
 40. Rahman S, Archana A, Dutta D, et al. The onus of cannabinoids in interrupting the molecular odyssey of breast cancer: A critical perspective on UPRER and beyond. *Saudi Pharm J* 2019;27:437-45.
 41. Vafaizadeh V, Buechel D, Rubinstein N, et al. The interactions of Bcl9/Bcl9L with β -catenin and Pygopus

- promote breast cancer growth, invasion, and metastasis. *Oncogene* 2021;40:6195-209.
42. Nath D, Li X, Mondragon C, et al. *Abi1* loss drives prostate tumorigenesis through activation of EMT and non-canonical WNT signaling. *Cell Commun Signal* 2019;17:120.
 43. Yu J, Zhou F, Luo D, et al. ZNF208 polymorphisms associated with ischemic stroke in a southern Chinese Han population. *J Gene Med* 2017.
 44. Lee E, Takita C, Wright JL, et al. Genome-wide enriched pathway analysis of acute post-radiotherapy pain in breast cancer patients: a prospective cohort study. *Hum Genomics* 2019;13:28.
 45. Dong H, Wang W, Mo S, et al. Long non-coding RNA SNHG14 induces trastuzumab resistance of breast cancer via regulating PABPC1 expression through H3K27 acetylation. *J Cell Mol Med* 2018;22:4935-47.
 46. Wu YQ, Ju CL, Wang BJ, et al. PABPC1L depletion inhibits proliferation and migration via blockage of AKT pathway in human colorectal cancer cells. *Oncol Lett* 2019;17:3439-45.
 47. Sasamoto N, Babic A, Rosner BA, et al. Development and validation of circulating CA125 prediction models in postmenopausal women. *J Ovarian Res* 2019;12:116.
 48. Fraschilla I, Pillai S. Viewing Siglecs through the lens of tumor immunology. *Immunol Rev* 2017;276:178-91.
 49. Chiarucci C, Cannito S, Daffinà MG, et al. Circulating Levels of PD-L1 in Mesothelioma Patients from the NIBIT-MESO-1 Study: Correlation with Survival. *Cancers (Basel)* 2020;12:361.
 50. Guo M, Yuan F, Qi F, et al. Expression and clinical significance of LAG-3, FGL1, PD-L1 and CD8+T cells in hepatocellular carcinoma using multiplex quantitative analysis. *J Transl Med* 2020;18:306.
 51. Fang WL, Chen MH, Huang KH, et al. The Clinicopathological Features and Genetic Alterations in Epstein-Barr Virus-Associated Gastric Cancer Patients after Curative Surgery. *Cancers (Basel)* 2020;12:1517.
 52. Newell F, Wilmott JS, Johansson PA, et al. Whole-genome sequencing of acral melanoma reveals genomic complexity and diversity. *Nat Commun* 2020;11:5259.
 53. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature* 2013;500:415-21.
 54. Teugels E, De Brakeleer S. An alternative model for (breast) cancer predisposition. *NPJ Breast Cancer* 2017;3:13.
 55. Alexandrov LB, Ju YS, Haase K, et al. Mutational signatures associated with tobacco smoking in human cancer. *Science* 2016;354:618-22.
 56. Wang S, Miller SR, Ober EA, et al. Making It New Again: Insight Into Liver Development, Regeneration, and Disease From Zebrafish Research. *Curr Top Dev Biol* 2017;124:161-95.
 57. Lim R, Sugino T, Nolte H, et al. Deubiquitinase USP10 regulates Notch signaling in the endothelium. *Science* 2019;364:188-93.
 58. Wang J, Zhang Y, Marian C, et al. Identification of aberrant pathways and network activities from high-throughput data. *Brief Bioinform* 2012;13:406-19.
 59. Deqin M, Chen Z, Nero C, et al. Somatic deletions of the polyA tract in the 3' untranslated region of epidermal growth factor receptor are common in microsatellite instability-high endometrial and colorectal carcinomas. *Arch Pathol Lab Med* 2012;136:510-6.
 60. Zhang Q, Cui Y, Zhang J, et al. Comparison of the characteristics of uncommon epidermal growth factor receptor (EGFR) mutations and EGFR-tyrosine kinase inhibitor treatment in patients with non-small cell lung cancer from different ethnic groups. *Exp Ther Med* 2020;19:3513-20.
 61. Yang JH, Li XY, Wang X, et al. Long-term persistent infection of HPV 16 E6 up-regulate SP1 and hTERT by inhibiting LKB1 in lung cancer cells. *PLoS One* 2017;12:e0182775.
 62. Wang Y, Guo J, Wang L, et al. Transcriptome analysis revealed potential mechanisms of differences in physiological stress responses between caged male and female magpies. *BMC Genomics* 2019;20:447.
 63. Barzilai N, Huffman DM, Muzumdar RH, et al. The critical role of metabolic pathways in aging. *Diabetes* 2012;61:1315-22.
 64. Yao JX, Chen X, Zhu YJ, et al. Prognostic Value of Vimentin Is Associated With Immunosuppression in Metastatic Renal Cell Carcinoma. *Front Oncol* 2020;10:1181.
 65. Chen Y, Zhang Y, Chai X, et al. Correlation between the Expression of PD-L1 and Clinicopathological Features in Patients with Thymic Epithelial Tumors. *Biomed Res Int* 2018;2018:5830547.
- (English Language Editor: J. Reynolds)

Cite this article as: Yang W, Chen S, Cheng X, Xu B, Zeng H, Zou J, Su C, Chen Z. Characteristics of genomic mutations and signaling pathway alterations in thymic epithelial tumors. *Ann Transl Med* 2021;9(22):1659. doi: 10.21037/atm-21-5182

Table S1 Patient clinical information

Group	ID	Age	Gender	Clinicopathologic diagnosis	Drinking history	Hypertension history	Metastatic sites	PD-L1
Thymic carcinoma	S01	67	F	Thymic carcinoma	No	No	None	Positive
	S04	47	M	Thymic carcinoma	Yes	No	None	Positive
	S05	39	M	Thymic carcinoma	No	No	None	Positive
	S06	42	M	Thymic carcinoma	No	No	Lung	Negative
	S10	69	M	Thymic carcinoma	No	Yes	None	Negative
	S12	58	M	Thymic carcinoma	Yes	No	None	Negative
	S13	65	F	Thymic carcinoma	No	Yes	None	Negative
	S14	57	M	Thymic carcinoma	No	Yes	None	Negative
	S15	52	M	Thymic carcinoma	No	No	None	Negative
Thymoma	S02	58	F	Thymoma	No	No	None	Positive
	S03	31	M	Thymoma	Yes	No	None	Positive
	S07	21	M	Thymoma	Yes	No	None	Negative
	S08	50	F	Thymoma	No	No	None	Negative
	S09	34	F	Thymoma	No	No	None	Negative
	S11	51	M	Thymoma	No	No	None	Negative
	S16	54	M	Thymoma	Yes	No	None	Positive
	S17	57	M	Thymoma	No	No	None	Negative
	S18	59	M	Thymoma	No	Yes	None	Positive
	S19	41	M	Thymoma	No	No	None	Positive
	S20	58	F	Thymoma	No	No	None	Positive
	S21	61	M	Thymoma	No	No	None	Positive