Am J Cancer Res 2025;15(3):1189-1200
www.ajcr.us /ISSN:2156-6976/ajcr0161899

Original Article

Distinct genomic features and mutational signatures of nucleotide excision repair and mismatch repair in thymoma

Po-Liang Cheng[1*], Wei-Jan Wang[2*], Cheng-Yen Chuang[3], Chih-Hung Lin[3], Chih-Yang Huang[4,5,6], Tzu-Hung Hsiao[1,7,8], Chung-Ping Hsu[3,9]

[1]Department of Medical Research, Taichung Veterans General Hospital, Taichung, Taiwan; [2]Department of Biological Science and Technology, China Medical University, Taichung, Taiwan; [3]Division of Thoracic Surgery, Department of Surgery, Taichung Veterans General Hospital, Taichung, Taiwan; [4]Cardiovascular and Mitochondrial Related Disease Research Center, Hualien Tzu Chi Hospital, Hualien, Taiwan; [5]Center of General Education, Buddhist Tzu Chi Medical Foundation, Tzu Chi University of Science and Technology, Hualien, Taiwan; [6]Department of Medical Research, China Medical University Hospital, China Medical University, Taichung, Taiwan; [7]Department of Public Health, Fu Jen Catholic University, New Taipei, Taiwan; [8]Institute of Genomics and Bioinformatics, National Chung Hsing University, Taichung, Taiwan; [9]Division of Thoracic Surgery, Department of Surgery, Buddhist Tzu Chi Medical Center, Hualien Tzu Chi Hospital, Hualien, Taiwan. [*]Equal contributors.

Received November 11, 2024; Accepted February 5, 2025; Epub March 15, 2025; Published March 30, 2025

Abstract: Thymoma is a rare malignancy with an unclear etiology of occurrence and development. We observed a higher incidence of thymoma in the Taiwanese population compared to other Western populations, suggesting the existence of different genomic features. Since most genomic studies are based on Western populations, we aimed to characterize the genomic profile of the Taiwanese population and compare it to the TCGA cohort in this study. We analyzed the genome of 47 thymoma patients using the Tumor Mutational Burden Panel to discover the genetic profile of the Taiwanese population. We also characterized the mutational signatures of these samples. Additionally, we leveraged RNA seq to estimate the gene expression profile and explored the featured pathways of thymoma in the Taiwanese population through gene set enrichment analysis.We identified several frequently mutated genes related to transcription, such as FAT1, KMT2D, and ZFHX3, as well as consensus mutational signatures associated with nucleotide excision repair (NER) and mismatch repair (MMR) deficiency. Our study also revealed increased activity of NER and MMR functions in our study cohort. Upon comparison with the TCGA cohort, we found dramatic differences in the most frequently mutated genes and mutational profiles between the Taiwanese and TCGA cohorts. Furthermore, we identified mismatch repair deficiency as a Taiwanese population-specific mutational signature with higher activity. These results highlight the distinct genomic background and molecular mechanisms of thymoma in the Taiwanese population, which may contribute to the development of new diagnostic and therapeutic strategies in the future.

Keywords: Thymoma, genomic features, mutational signatures, DNA repair mechanisms, population-specific differences

Introduction

Thymoma is a rare type of thymic epithelial tumor that consists of varying proportions of epithelial and lymphocytic components. The World Health Organization (WHO) classifies thymic epithelial tumors (TETs) as thymoma (types A, AB, B1, B2, and B3) or thymic carcinoma based on their morphological features and the relative proportion of lymphocytes [1]. However, effective treatments for thymoma are still lacking. Surgical resection remains the primary approach, and the prognosis is influenced by the tumor stage at diagnosis and the extent of surgical removal [2, 3]. Despite the existence of various other treatment options, none of them are curative. The incidence of thymoma in western countries is 0.13-0.15 per 100,000 population at risk [4, 5]. In the America, Asian/Pacific Islanders have a higher incidence of thymoma

https://doi.org/10.62347/ZWSB8391

than other racial groups [6]. Similarly, a higher incidence of thymoma has been observed in Taiwanese populations. According to the Cancer Registry Annual Report 2016 in Taiwan, the incidence of thymic cancer is 1.61 per 100,000 population at risk in men and 1.33 per 100,000 population at risk in women [6]. Thus, this recent study suggested the presence of different genomic risk factors in the Taiwanese population. However, there is still a lack of relevant research at present.

Thymoma is commonly diagnosed based on its morphology, various studies have offered insights into the immunohistochemical characteristics of these tumors. Presently, there is no established unique biomarker for thymic epithelial tumor or thymoma. However, recent studies have indicated that epithelial cells in thymic epithelial neoplasms, including thymoma, may express polyclonal PAX8 and PD-L1 [7, 8]. These studies showed that high PD-L1 expression was associated with advanced staging and high-grade histology. Over the past decade, numerous investigations have aimed to characterize the molecular profile of thymoma. Although some findings support the need for targeted biomarker treatment for thymoma, the genomic background of thymoma remains largely unknown. Previous studies have demonstrated certain proteins in the immunohistochemical characteristics of thymoma. For instance, a previous study using exome sequencing revealed a high frequency of recurrent mutations in the GTF2I gene in thymoma [9]. Additionally, several genes, including NRAS, HRAS, and TP53, have been identified as mutated oncogenes. Moreover, immunohistochemical analysis has shown the overexpression of EGFR, HER2, KIT, and IGF-1R in thymoma, while the activation of β-catenin has been linked to thymoma initiation and progression [10, 11]. These studies have led to the identification of a new molecular subtyping of thymic epithelial tumors (TETs) associated with disease-free and overall survival [12]. However, most of these studies have been conducted on Western populations, and the genomic profile of thymoma in the Taiwanese population remains unclear. Additionally, there are currently no well-established unique markers for thymoma in Taiwan.

In this study, we aimed to characterize the distinct genomic features between the Taiwanese and Western populations using a tumor mutational target panel and the TCGA database. And our findings revealed specific frequently mutated genes and mutational profiles in the Taiwanese population, which will enhance our understanding of the biological features of thymoma and address clinical needs.

Materials and methods

Samples

From April 20, 2006 to December 29, 2020, 47 thymoma patients who received surgical resection were retrospectively enrolled in this study. We get the specimen from the BioBank of Taichung Veterans General Hospital (TVGH) and apply to perform NGS. We collected clinical information such as age, sex, World Health Organization histologic classification, and history of myasthenia gravis. All subjects provided written informed consent, and the study protocol was approved by the Institutional Review Board (IRB) of TVGH (SF16054B-1).

The genomic analysis of tumor tissue and data collection were approved by the Institutional Review Board of Taichung Veterans General Hospital. DNA of 47 patients with thymoma specimens was extracted using the DNeasy Blood & Tissue Kit (Qiagen GmbH, QIAGEN Strasse 1, 40724 Hilden, Germany), according to the manufacturer's instructions.

Target panel sequencing and genomic analysis

DNA libraries were generated by QIAseq Tumor Mutational Burden Panel (), which covers 486 genes (1.3 Mbp), with 40 ng of tumor genomic DNA from all samples following the manufacturer's instruction. The prepared library was sequenced with paired-end runs by Nextseq 550 sequencer. DNA reads were mapped to the human reference genome GRCh38 and variant calling by built workflow in CLC software. Somatic mutation annotation filtering and characterization were performed by QIAGEN Clinical Insight (QCI). After removing variants which is reported in Taiwan biobank (TWB) frequency ≥ 1% or is reported as benign for diseases by ClinVar, pathogenic and likely patho-

Table 1. Characteristics of the enrolled thymoma patients

		Total (%)
Total Number		47
Age (years), median (range)		46.96 (24-80)
Gender	Male	16 (34.0%)
	Female	31 (66.0%)
WHO Type	WHO type A	3 (6.4%)
	WHO type AB	14 (29.8%)
	WHO type B1	10 (21.3%)
	WHO type B2	14 (29.8%)
	WHO type B3	6 (12.8%)
History of myasthenia gravis		20 (43%)

genic variants were identified according to ACMG variant interpretation guidelines [13].

Mutational signature analysis

Mutational signatures in thymoma samples were analyzed by R package MutationalPatterns [14]. Single nucleotide variants identified from target panel sequencing was converted into 6 types of substitutions (C > A, C > G, C > T, T > A, T > C, T > G). Then, quantifying the contribution of COSMIC signatures in each thymoma sample by fitting constructed 96 trinucleotide changes matrix in each sample to COSMIC mutational signatures (https://cancer.sanger.ac.uk/signatures/). Mutational signature contribution of thymoma in TCGA were download from previous study [15].

Gene expression analysis

Sequence reads were mapped to the human reference genome GRCh38 by HISAT2 [16]. FeatureCounts was used to calculate read counts of each expressed gene [17], and the differentially expressed genes between thymoma tissues and peripheral normal tissues were identified by R package DEseq2 [18]. When the adjusted *p*-value of DGE < 0.05 was considered as statistical significant. The GEPIA2 database (http://gepia2.cancer-pku.cn/#analysis) was utilized to analyze gene expression differences.

Gene set enrichment analysis

Total genes from RNAseq analysis were pre-ranked according to signed -log10 *P* values based on the significance of differential gene expression between thymoma and peripheral normal tissues. Gene set C2 from the Molecular Signatures Database (MSigDB) was used to perform enrichment analysis on ranked genes. Only when the nominal *p*-value < 0.05 was considered as statistically significant.

Results

The frequently mutated genes across thymoma patients in the Taiwanese population

To identify the potential biomarkers in thymoma, we conducted targeted panel sequencing on 47 thymoma samples, encompassing histological subtypes according to the WHO classification (A, AB, B1, B2, and B3), with the corresponding clinical characteristics detailed in **Table 1**. In total, we identified 347 mutations, including 23 indels, across all thymoma samples. In **Figure 1A**, we identified the genes with pathogenic or likely pathogenic mutations occurring in more than 10% of cases. Notably, FAT1 was the most frequently mutated gene, observed in 21% of all thymoma samples. Moreover, we also found other frequently mutated genes included ZFHX3 (12.7%), KMT2D (12.7%), LRP1B (14.8%), MSH6 (10.6%), KEL (10.6%), and LIG3 (10.6%).

Furthermore, we surprisingly identified several age-related mutations. For instance, patients with MSH6, AXL, CD276, or SPEN mutations had an average age of 34.0±7.9, 35.0±4.3, 35.6±4.9, and 36.6±2.08 years, respectively, significantly lower than the average age of patients without these mutations (48.5±12.1, 47.7±12.5, 47.7±12.5, and 47.6±12.6, respectively, all *p*-value < 0.05) (**Figure 1B**). Conversely, patients with VSIR, DNMT3A, and PDIA3 mutations had an average age of 58.5±2.1, 66±1.4, and 64.5±3.5 years, respectively, significantly higher than the average age of patients without these mutations (46.4±12.5, 46.1±12.1, and 46.1±12.2, respectively, all *p*-value < 0.05) (**Figure 1B**). However, we did not find any single variant correlated with myasthenia gravis (MG) or the WHO classification. In sum, we found frequent mutations in genes such as FAT1, ZFHX3, KMT2D, LRP1B, MSH6, KEL, and LIG3, while also identifying age-related mutation patterns in genes like MSH6, AXL, CD276, SPEN, VSIR, DNMT3A, and PDIA3 in samples from thymoma patients.

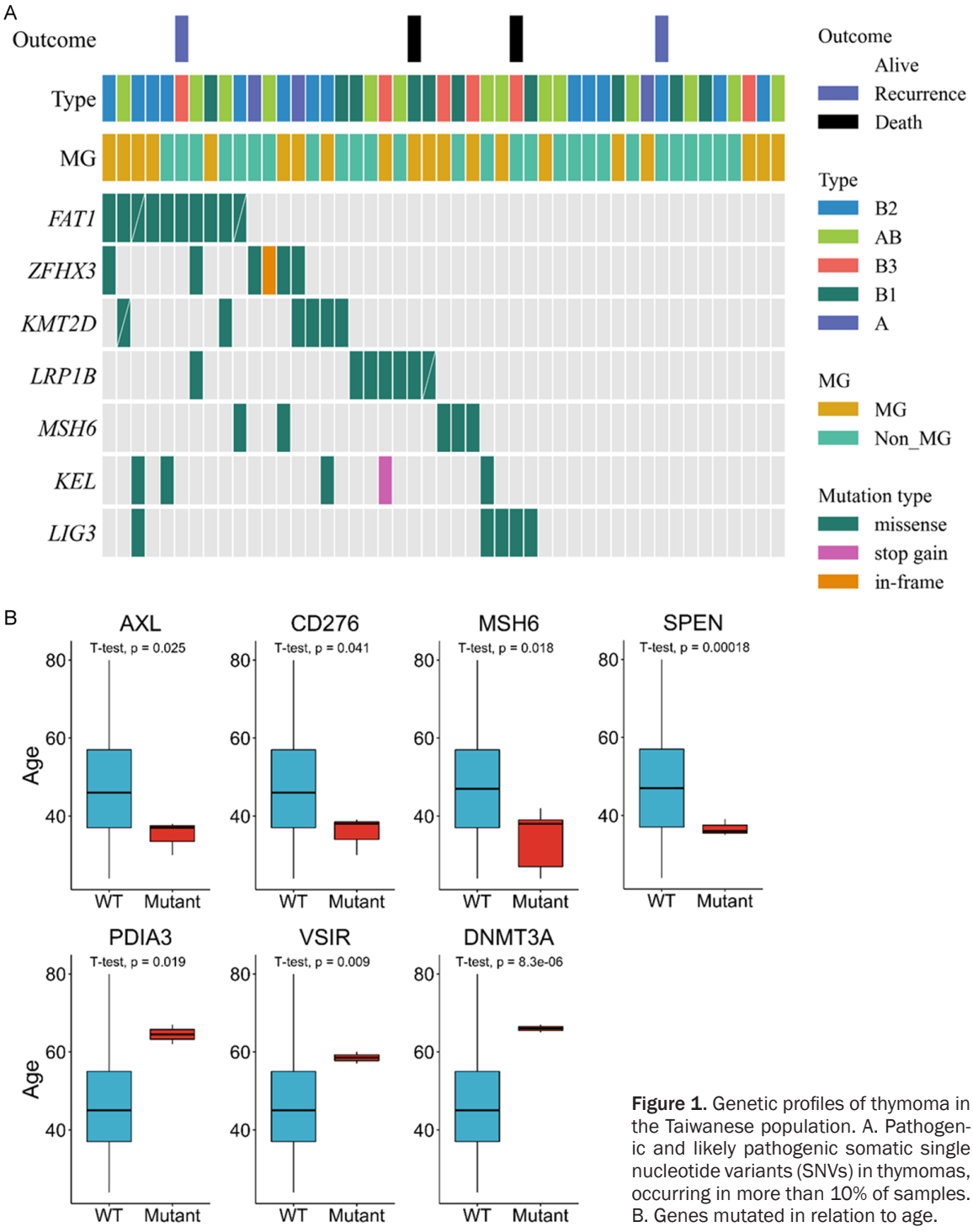


Figure 1. Genetic profiles of thymoma in the Taiwanese population. A. Pathogenic and likely pathogenic somatic single nucleotide variants (SNVs) in thymomas, occurring in more than 10% of samples. B. Genes mutated in relation to age.

Consistent mutational profiles of thymoma in the Taiwanese population

As previously reported, SBS5, also referred to as Signature 5, is a mutational signature identified in cancer genomes through the analysis of somatic mutation patterns. SBS5 has been associated with aging, tobacco smoking, and an unknown deficiency in nucleotide excision repair (NER), whereas SBS26 is specifically linked to deficiencies in mismatch repair (MMR). [19]. These mutational signatures consistently contributed to all thymoma samples. Thus, to further identify cancer genomes, we used

mutational signature analysis and found two significant contributors related to previously described signatures, SBS5 and SBS26 (**Figure 2A**).

Given the multiple possible causes of SBS5, we investigated its correlation with age and smoking behavior. However, we found no correlation between the contribution of SBS5 and age (Pearson's correlation, r = -0.14, t = -0.92, df = 45, *p*-value = 0.36) or any difference between patients with and without a history of smoking (t-test, *p*-value = 0.48). Additionally, the contribution of the SBS5 signature showed no correlation with age in patients with age-related mutations (Pearson's correlation, r = -0.06, t = -0.22, df = 14, *p*-value = 0.83) (**Figure 2A**). Therefore, NER deficiency might be the potential cause of the SBS5 signature in thymoma patients in the Taiwanese population.

Deregulation of DNA repair is associated with the mutational profiles of thymoma

Furthermore, we observed that the mutational signatures related to DNA repair processes, such as SBS5 and SBS26, accounted for more than 50% of the total mutational signatures in each sample. To further investigate the impact of dysregulated DNA repair in thymoma, we analyzed somatic mutations in 450 DNA damage response-associated genes (DDR-associated genes) curated by experts [20, 21]. In our study cohort, we identified 76 somatic SNVs and indels in 46 DDR-associated genes. Among these, POLE (17% of samples), MSH6 (10.6% of samples), and LIG3 (10.6% of samples) were frequently mutated genes associated with nucleotide excision repair (NER) and mismatch repair (MMR) (**Figure 2B**). Through gene set enrichment analysis (GSEA), we found that NER- and MMR-associated genes were enriched in tumor tissues compared to peripheral normal tissues (**Figure 2C**, **2D**; NER: ES = 0.55, P < 0.001; MMR: ES = 0.75, P = 0.018). Additionally, we observed elevated expression of ERCC2 and RPA1, which are part of the leading-edge subset of NER and MMR gene sets, in mutated thymoma samples (**Figure 2B**).

Distinct genomic characteristics between the Taiwanese and western populations

The distribution of histological classifications in the TCGA cohort was similar to our study cohort (chi-squared test, P = 0.46), with 27% of patients (n = 34) diagnosed with MG. In addition, upon comparing the frequently mutated genes between the TCGA and Taiwanese populations, we found that the most commonly mutated gene in the TCGA cohort was GTF2I (33% of samples), followed by HRAS (8% of samples) and TP53 (3% of samples), which were not detected in the Taiwanese population (**Figure 3A**). Conversely, frequently mutated genes in the Taiwanese population, such as FAT1 (21% of samples), ZFHX3 (12.7% of samples), and KMT2D (12.7% of samples), were either not detected or found in fewer instances in the TCGA cohort (0%, 1.6%, and 0%, respectively). Furthermore, although GTF2I was not detected by our tumor mutational burden target panel, we attempted to identify variants by using RNAseq data. However, only 11 variants were detected, and upon further analysis via Sanger sequencing, we confirmed that the identified nonsynonymous mutation was not a genuine variant, indicating that GTF2I variants were not present in our study cohort. Additionally, we did not detect HRAS, with only one NRAS mutant identified in our study cohort.

The differences of mutational signatures were found between the Taiwanese and TCGA cohorts, including SBS1 and SBS5. In the Taiwanese cohort, SBS1 and SBS5 contributed 8.8±1.9% and 28.4±4.7%, respectively, whereas in the TCGA cohort, they contributed 11.9±10.8% and 41.3±30.5%, respectively (**Figure 3B**). Instead, we identified SBS26 as a Taiwanese-specific mutational signature, contributing 22.9±2.4% and not detected in the TCGA cohort (**Figure 3B**). Through gene set enrichment analysis (GSEA), we observed that DNA repair-related functions had a higher enrichment score in the Taiwanese cohort than in the TCGA cohort, particularly in MMR (TW: ES = 0.74, P < 0.05; TCGA: ES = 0.52, P = 0.17) (**Figure 3C** and Supplementary Figure 1). Furthermore, to verify the role of NER and MMR in thymoma, we analyzed ERCC2 gene expression using the GEPIA2 database based on TCGA data. The results indicated that ERCC2 is highly expressed in thymoma tumor tissue, suggesting that NER and MMR may play a critical role in thymoma progression (**Figure 3D**). Taken together, our results demonstrated that the differences in frequently mutated

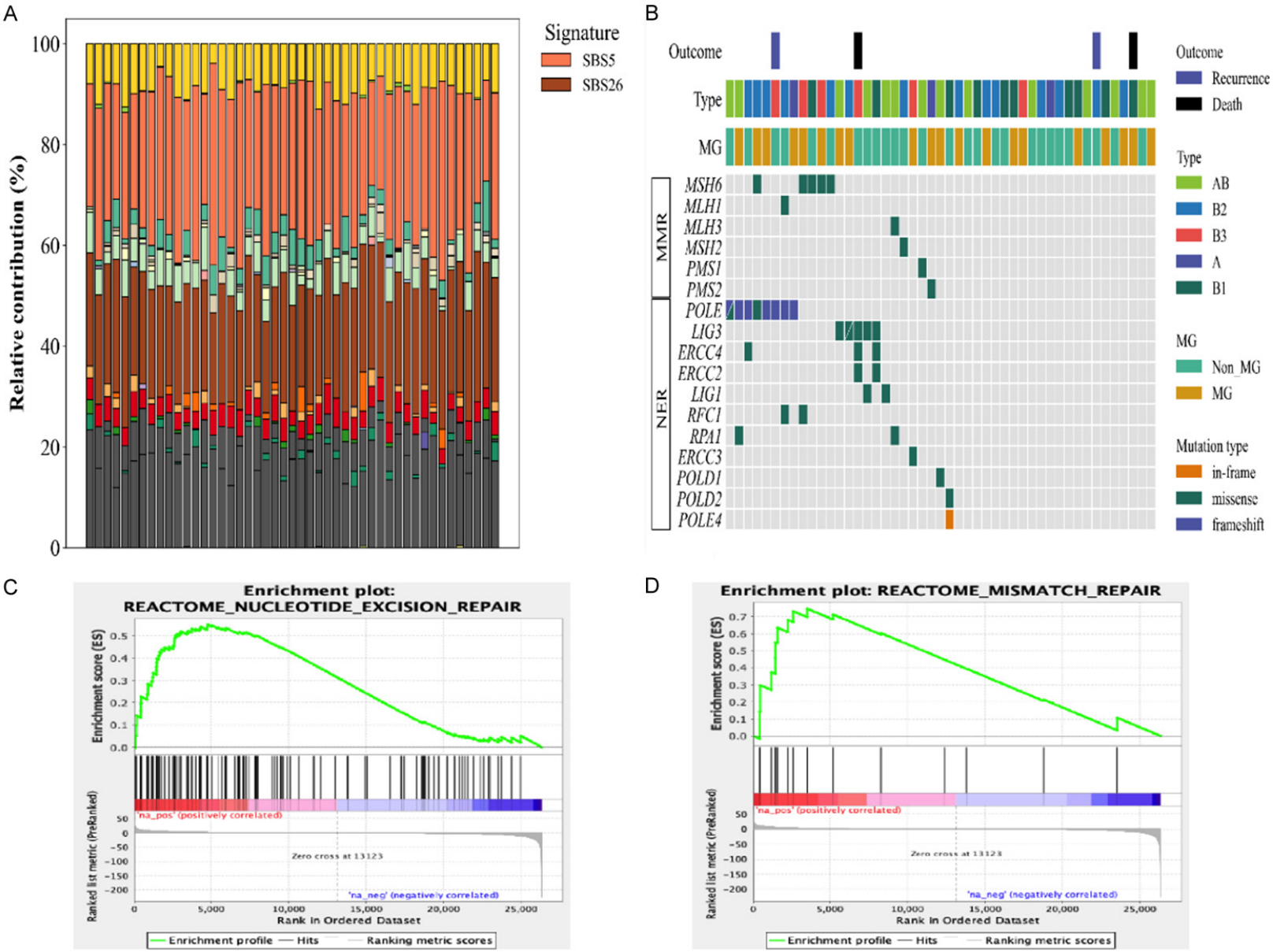


Figure 2. Deregulation of DNA repair genes is associated with the mutational profiles of thymoma. A. Consensus mutational profile among all thymoma patients. B. Single nucleotide variants (SNVs) and Indels of nucleotide excision repair (NER) and mismatch repair (MMR)-related genes in thymoma patients. C and D. Gene set enrichment analysis (GSEA) of NER and MMR-associated genes conducted in thymomas compared with normal tissues.

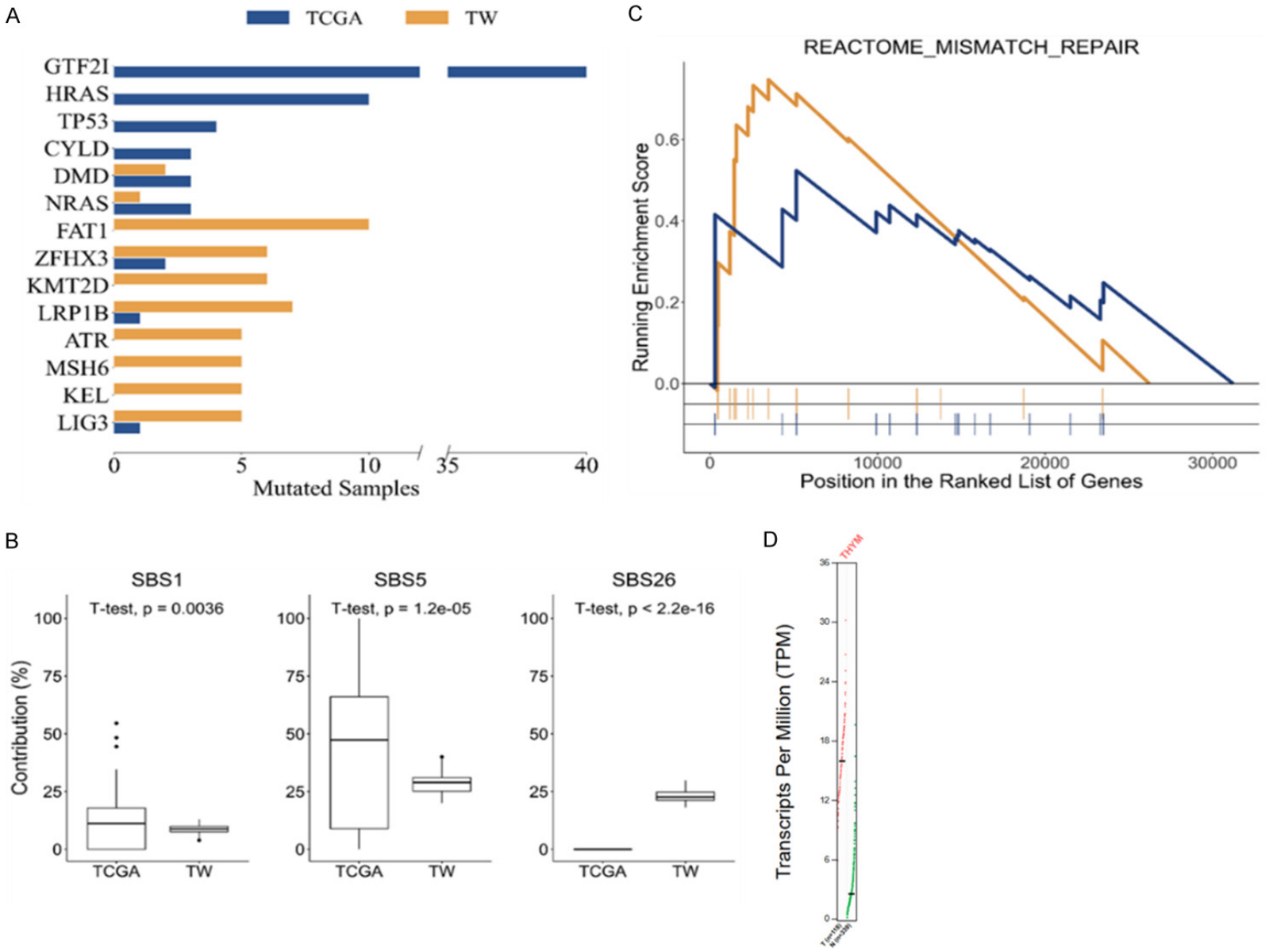


Figure 3. Depicts the distinct genomic features observed in the TCGA cohort compared to the Taiwanese population. A. The frequently mutated genes of thymoma in the TCGA cohort compare to the Taiwanese population. B. The most contributed mutational signatures in the TCGA compares to Taiwanese populations. C. The activity of the MMR pathway in the TCGA compares to Taiwanese populations by using GSEA. D. The expression of the ERCC2 gene in thymoma from the TCGA dataset.

genes and mutational signatures between the TCGA and Taiwanese populations.

Discussion

In this study, we used multiple sequencing tools to identify several novel transcription-related genes that are frequently mutated including FAT1, KMT2D, and ZFHX3 in thymoma tissue samples. Among these, FAT1, the most frequently mutated gene, is a tumor suppressor gene and has been demonstrated that inactivation of FAT1 promotes Wnt pathway signaling activation and cancer cell proliferation [20, 22]. In consistence with our findings, previous study identified that activation of Wnt pathway is a molecular subtype of thymoma by genomic analysis [12]. Another frequently mutated gene, the histone methyltransferase KMT2D, plays critical roles in regulating cell fate transition [23], metabolism [24] and tumor suppression [25]. Moreover, KMT2D deficiency impairs the regulation of the glycolysis program and promotes lung tumorigenesis [26]. However, the role of ZFHX3 in carcinogenesis is controversial. Previous studies indicate that ZFHX3 is crucial for angiogenesis in liver cancer cells [27] and is reported to promote breast cancer proliferation and growth [28]. On the contrary, ZFHX3 is also identified as a tumor suppressor, and its defects are associated with a poor outcome in endometrial cancer [29]. In our study cohort, we did not find overexpression of ZFHX3 in thymoma compared to normal peripheral tissues. Therefore, we suggest ZFHX3 may play a tumor suppressor role in thymoma. Overall, these findings suggest that the frequent mutation of these transcription-related genes may be involved in the carcinogenesis of thymoma in the Taiwanese population.

In addition, our previous study identified 140 differentially expressed genes associated with myasthenia gravis (MG) in thymoma patients. Among these genes, we observed significant upregulation of hypoxia-inducible factor 3 alpha (HIF3A), insulin-like growth factor-binding protein 1 (IGFBP1), pyruvate dehydrogenase kinase (PDK) and Krüppel-like factor 15 (KLF15) in patients with thymoma who had MG compared to those without MG [30]. These genes show promise as potential biomarkers for predicting the development of MG in thymoma patients or as targets for therapeutic interventions. However, the detailed genetic mutations of thymoma remain largely unknown. In this study, we discovered that FAT1, KMT2D and ZFHX3 are highly mutated in the Taiwanese population and in the TCGA database of thymoma. However, we did not observe a statistically significant correlation with MG. Taken together, thymomas with or without MG may involve different underlying genetic regulatory mechanisms, such as epigenetic regulation, which warrant further investigation.

A previous study found that increased MMR-associated genes ERCC2 expression is correlated with the tumor stage and grade of head and neck cancer [31] and downregulation of RPA1 enhances the radiosensitivity of nasopharyngeal cancer [32]. Moreover, the NER and MMR functions are well known to be activated in tumor tissues compared to peripheral normal tissues. Thus, it suggests that ERCC2 and RPA1 might be involved in the carcinogenesis and chemoresistance of thymoma in the Taiwanese population. In this study, we also observed that SBS 5 and SBS26 were the predominant mutational signatures, comprising more than 50% of the total mutational profile among all samples in the Taiwanese population. Since these two signatures are related to NER and MMR, respectively, this might suggest a strong mutagenic effect of DNA repair dysfunction in thymoma patients. Following these findings, we showed that NER and MMR related genes ERCC2 and RPA1, are found to be overexpressed and mutated in thymoma tissues in our study cohort. Furthermore, previous research has shown that increased ERCC2 expression is associated with aggressive tumors in head and neck cancer [31], while reduced expression of MMR genes, such as MSH2/MSH6, promotes pituitary tumor growth [33].

Thus, given the increased activity of NER and MMR pathways in thymoma, targeting these DNA repair mechanisms presents a promising therapeutic strategy. Notably, FDA-approved drugs such as immune checkpoint inhibitors (e.g., pembrolizumab and dostarlimab) have shown efficacy in tumors with MMR deficiencies [34], while agents like PARP inhibitors (e.g., olaparib) indirectly impact the NER pathway by exploiting synthetic lethality [35], suggesting potential avenues to enhance chemo-

therapeutic response in thymoma patients. According to our results, these mutated genes may disrupt the DNA repair process, leading to the NER and MMR-related mutational signatures observed in thymoma patients. However, MMR is the last line of defense when a misincorporation fails in proofreading [36]. Therefore, impaired MMR genes result in an increased mutation rate in cancer cells [37]. Moreover, we identified mutational signature SBS26, which is described as MMR deficiency, as a unique signature in the Taiwanese population. The increased DNA repair capacity, in coordination with the impaired DNA repair function, may cause the special genetic profile observed in thymomas in the Taiwanese population.

GTF2I is the most frequently mutated gene in thymoma and is reported as a marker of favorable prognosis [9, 38]. However, the Taiwanese population did not show a lower survival rate when lacking GTF2I mutants. Mutations of the TP53, HRAS, and NRAS genes are also frequently discovered in thymoma in the TCGA database [12, 39]. HRAS and NRAS are GTPases involved in cell proliferation and survival [40] and are mainly found in type A and AB thymomas [41]. In addition, TP53 is a well-known tumor suppressor gene, and its expression might be a prognostic factor in thymoma [42]. However, thymomas in the Taiwanese population also lack mutations of TP53. Therefore, these genes may not be involved in the onset of thymoma in the Taiwanese population. Furthermore, impaired DNA repair pathways can increase genomic instability. In this study, we found that DNA repair processes were activated in the tumor tissues of both the Taiwanese and TCGA cohorts, but most DNA repair pathways, especially MMR, had higher activity in the Taiwanese population than in the TCGA cohort. Taken together, our study used the limited sequencing method. By using Target panel sequencing has a limited scope compared to whole-exome sequencing (WES) or whole-genome sequencing (WGS) coverage. Therefore, we may have missed some differences in regions not covered by the target panel, even though the panel comprised most tumor driver genes. In addition, while a target panel can be used to detect mutational profiles [43], there may be some bias in identifying mutational signatures due to the composition of the trinucleotide context not being as comprehensive as initially described for these signatures.

Conclusion

In this study, we present a comprehensive genomic characterization of thymoma in the Taiwanese population, highlighting its novel aspects. Through comparative analysis with the TCGA cohort, we identified distinct differences in the frequently mutated genes FAT1, KMT2D, and ZFHX3, as well as unique mutational signatures specific to the Taiwanese population. Moreover, our findings suggest that differential activation of DNA repair pathways, along with mutations in nucleotide excision repair (NER) and mismatch repair (MMR) genes, may contribute to the observed genomic variations. These insights provide a foundation for further investigation and validation of these unique genomic features as potential biomarkers for prognostic assessment and therapeutic targeting in Taiwanese patients with thymoma.

Acknowledgements

This research was funded by grants from the Taiwanese National Science and Technology Council (grant numbers: NSTC 112-2320-B-039-012-MY3 to W-J. W. and MOST 105-2314-B-075A-008, MOST 106-2314-B-075A-006-MY2 to T-H. H.). This research was also funded by grants from China Medical University (CMU113-MF-69 to W-J. W.).

Disclosure of conflict of interest

None.

Address correspondence to: Chung-Ping Hsu, Division of Thoracic Surgery, Department of Surgery, Taichung Veterans General Hospital, Taichung, Taiwan. E-mail: cliffhsu@gmail.com; Tzu-Hung Hsiao, Department of Medical Research, Taichung Veterans General Hospital, Taichung, Taiwan. E-mail: thsiao@vghtc.gov.tw

References

[1] Kumar R. Myasthenia gravis and thymic neoplasms: a brief review. World J Clin Cases 2015; 3: 980-983.
[2] Detterbeck FC and Parsons AM. Thymic tumors. Ann Thorac Surg 2004; 77: 1860-1869.
[3] Conforti F, Zucali PA, Pala L, Catania C, Bagnardi V, Sala I, Della Vigna P, Perrino M, Zagami

P, Corti C, Stucchi S, Barberis M, Guerini-Rocco E, Di Venosa B, De Vincenzo F, Cordua N, Santoro A, Giaccone G and De Pas TM. Avelumab plus axitinib in unresectable or metastatic type B3 thymomas and thymic carcinomas (CAVEATT): a single-arm, multicentre, phase 2 trial. Lancet Oncol 2022; 23: 1287-1296.

[4] Engels EA and Pfeiffer RM. Malignant thymoma in the United States: demographic patterns in incidence and associations with subsequent malignancies. Int J Cancer 2003; 105: 546-551.

[5] Engels EA. Epidemiology of thymoma and associated malignancies. J Thorac Oncol 2010; 5 Suppl 4: S260-265.

[6] Hsu CH, Chan JK, Yin CH, Lee CC, Chern CU and Liao CI. Trends in the incidence of thymoma, thymic carcinoma, and thymic neuroendocrine tumor in the United States. PLoS One 2019; 14: e0227197.

[7] Weissferdt A and Moran CA. Pax8 expression in thymic epithelial neoplasms: an immunohistochemical analysis. Am J Surg Pathol 2011; 35: 1305-1310.

[8] Weissferdt A, Fujimoto J, Kalhor N, Rodriguez J, Bassett R, Wistuba II and Moran CA. Expression of PD-1 and PD-L1 in thymic epithelial neoplasms. Mod Pathol 2017; 30: 826-833.

[9] Petrini I, Meltzer PS, Kim IK, Lucchi M, Park KS, Fontanini G, Gao J, Zucali PA, Calabrese F, Favaretto A, Rea F, Rodriguez-Canales J, Walker RL, Pineda M, Zhu YJ, Lau C, Killian KJ, Bilke S, Voeller D, Dakshanamurthy S, Wang Y and Giaccone G. A specific missense mutation in GTF2I occurs at high frequency in thymic epithelial tumors. Nat Genet 2014; 46: 844-849.

[10] Scorsetti M, Leo F, Trama A, D'Angelillo R, Serpico D, Macerelli M, Zucali P, Gatta G and Garassino MC. Thymoma and thymic carcinomas. Crit Rev Oncol Hematol 2016; 99: 332-350.

[11] Liang CC, Lu TL, Yu YR, You LR and Chen CM. beta-catenin activation drives thymoma initiation and progression in mice. Oncotarget 2015; 6: 13978-13993.

[12] Lee HS, Jang HJ, Shah R, Yoon D, Hamaji M, Wald O, Lee JS, Sugarbaker DJ and Burt BM. Genomic analysis of thymic epithelial tumors identifies novel subtypes associated with distinct clinical features. Clin Cancer Res 2017; 23: 4855-4864.

[13] Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K and Rehm HL; ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015; 17: 405-424.

[14] Blokzijl F, Janssen R, van Boxtel R and Cuppen E. MutationalPatterns: comprehensive genome-wide analysis of mutational processes. Genome Med 2018; 10: 33.

[15] Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, Boot A, Covington KR, Gordenin DA, Bergstrom EN, Islam SMA, Lopez-Bigas N, Klimczak LJ, McPherson JR, Morganella S, Sabarinathan R, Wheeler DA, Mustonen V; PCAWG Mutational Signatures Working Group, Getz G, Rozen SG and Stratton MR; PCAWG Consortium. The repertoire of mutational signatures in human cancer. Nature 2020; 578: 94-101.

[16] Kim D, Langmead B and Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods 2015; 12: 357-360.

[17] Liao Y, Smyth GK and Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 2014; 30: 923-930.

[18] Anders S and Huber W. Differential expression analysis for sequence count data. Genome Biol 2010; 11: R106.

[19] Meier B, Volkova NV, Hong Y, Schofield P, Campbell PJ, Gerstung M and Gartner A. Mutational signatures of DNA mismatch repair deficiency in C. elegans and human cancers. Genome Res 2018; 28: 666-675.

[20] Morris LG, Kaufman AM, Gong Y, Ramaswami D, Walsh LA, Turcan S, Eng S, Kannan K, Zou Y, Peng L, Banuchi VE, Paty P, Zeng Z, Vakiani E, Solit D, Singh B, Ganly I, Liau L, Cloughesy TC, Mischel PS, Mellinghoff IK and Chan TA. Recurrent somatic mutation of FAT1 in multiple human cancers leads to aberrant Wnt activation. Nat Genet 2013; 45: 253-261.

[21] Chang LC and Yu YL. Dietary components as epigenetic-regulating agents against cancer. Biomedicine (Taipei) 2016; 6: 2.

[22] Chen M, Sun X, Wang Y, Ling K, Chen C, Cai X, Liang X and Liang Z. FAT1 inhibits the proliferation and metastasis of cervical cancer cells by binding beta-catenin. Int J Clin Exp Pathol 2019; 12: 3807-3818.

[23] Wang C, Lee JE, Lai B, Macfarlan TS, Xu S, Zhuang L, Liu C, Peng W and Ge K. Enhancer priming by H3K4 methyltransferase MLL4 controls cell fate transition. Proc Natl Acad Sci U S A 2016; 113: 11871-11876.

[24] Kim DH, Rhee JC, Yeo S, Shen R, Lee SK, Lee JW and Lee S. Crucial roles of mixed-lineage leukemia 3 and 4 as epigenetic switches of the hepatic circadian clock controlling bile acid homeostasis in mice. Hepatology 2015; 61: 1012-1023.

[25] Ortega-Molina A, Boss IW, Canela A, Pan H, Jiang Y, Zhao C, Jiang M, Hu D, Agirre X, Niesvizky I, Lee JE, Chen HT, Ennishi D, Scott DW, Mottok A, Hother C, Liu S, Cao XJ, Tam W, Shaknovich R, Garcia BA, Gascoyne RD, Ge K, Shilatifard A, Elemento O, Nussenzweig A, Melnick AM and Wendel HG. The histone lysine methyltransferase KMT2D sustains a gene expression program that represses B cell lymphoma development. Nat Med 2015; 21: 1199-1208.

[26] Alam H, Tang M, Maitituoheti M, Dhar SS, Kumar M, Han CY, Ambati CR, Amin SB, Gu B, Chen TY, Lin YH, Chen J, Muller FL, Putluri N, Flores ER, DeMayo FJ, Baseler L, Rai K and Lee MG. KMT2D deficiency impairs super-enhancers to confer a glycolytic vulnerability in lung cancer. Cancer Cell 2020; 37: 599-617, e597.

[27] Fu C, An N, Liu J, A J, Zhang B, Liu M, Zhang Z, Fu L, Tian X, Wang D and Dong JT. The transcription factor ZFHX3 is crucial for the angiogenic function of hypoxia-inducible factor 1alpha in liver cancer cells. J Biol Chem 2020; 295: 7060-7074.

[28] Dong G, Ma G, Wu R, Liu J, Liu M, Gao A, Li X, A J, Liu X, Zhang Z, Zhang B, Fu L and Dong JT. ZFHX3 promotes the proliferation and tumor growth of er-positive breast cancer cells likely by enhancing stem-like features and MYC and TBX3 transcription. Cancers (Basel) 2020; 12: 3415.

[29] Walker CJ, Miranda MA, O'Hern MJ, McElroy JP, Coombes KR, Bundschuh R, Cohn DE, Mutch DG and Goodfellow PJ. Patterns of CTCF and ZFHX3 mutation and associated outcomes in endometrial cancer. J Natl Cancer Inst 2015; 107: djv249.

[30] Lee MC, Hsiao TH, Chuang HN, Lee LW, Chi PL, Tsai HM, Mao CL and Hsu CP. Molecular profiling of thymoma with myasthenia gravis: risk factors of developing myasthenia gravis in thymoma patients. Lung Cancer 2020; 139: 157-164.

[31] Zafeer M, Mahjabeen I and Kayani MA. Increased expression of ERCC2 gene in head and neck cancer is associated with aggressive tumors: a systematic review and case-control study. Int J Biol Markers 2016; 31: e17-25.

[32] Zhang Z, Huo H, Liao K, Wang Z, Gong Z, Li Y, Liu C and Hu G. RPA1 downregulation enhances nasopharyngeal cancer radiosensitivity via blocking RAD51 to the DNA damage site. Exp Cell Res 2018; 371: 330-341.

[33] Uraki S, Ariyasu H, Doi A, Kawai S, Takeshima K, Morita S, Fukai J, Fujita K, Furuta H, Nishi M, Sugano K, Inoshita N, Nakao N, Yamada S and Akamizu T. Reduced expression of mismatch repair genes MSH6/MSH2 directly promotes pituitary tumor growth via the ATR-Chk1 pathway. J Clin Endocrinol Metab 2018; 103: 1171-1179.

[34] Abida W, Cheng ML, Armenia J, Middha S, Autio KA, Vargas HA, Rathkopf D, Morris MJ, Danila DC, Slovin SF, Carbone E, Barnett ES, Hullings M, Hechtman JF, Zehir A, Shia J, Jonsson P, Stadler ZK, Srinivasan P, Laudone VP, Reuter V, Wolchok JD, Socci ND, Taylor BS, Berger MF, Kantoff PW, Sawyers CL, Schultz N, Solit DB, Gopalan A and Scher HI. Analysis of the prevalence of microsatellite instability in prostate cancer and response to immune checkpoint blockade. JAMA Oncol 2019; 5: 471-478.

[35] Prosz A, Sahgal P, Huffman BM, Sztupinszki Z, Morris CX, Chen D, Borcsok J, Diossy M, Tisza V, Spisak S, Likasitwatanakul P, Rusz O, Csabai I, Cecchini M, Baca Y, Elliott A, Enzinger P, Singh H, Ubellaker J, Lazaro JB, Cleary JM, Szallasi Z and Sethi NS. Mutational signature-based identification of DNA repair deficient gastroesophageal adenocarcinomas for therapeutic targeting. NPJ Precis Oncol 2024; 8: 87.

[36] Ijsselsteijn R, Jansen JG and de Wind N. DNA mismatch repair-dependent DNA damage responses and cancer. DNA Repair (Amst) 2020; 93: 102923.

[37] Zhang Y, Yuan F, Presnell SR, Tian K, Gao Y, Tomkinson AE, Gu L and Li GM. Reconstitution of 5'-directed human mismatch repair in a purified system. Cell 2005; 122: 693-705.

[38] Lang M, Kazdal D, Mohr I and Anamaterou C. Differences and similarities of GTF2I mutated thymomas in different Eurasian ethnic groups. Transl Lung Cancer Res 2023; 12: 1842-1844.

[39] Radovich M, Pickering CR, Felau I, Ha G, Zhang H, Jo H, Hoadley KA, Anur P, Zhang J, McLellan M, Bowlby R, Matthew T, Danilova L, Hegde AM, Kim J, Leiserson MDM, Sethi G, Lu C, Ryan M, Su X, Cherniack AD, Robertson G, Akbani R, Spellman P, Weinstein JN, Hayes DN, Raphael B, Lichtenberg T, Leraas K, Zenklusen JC; Cancer Genome Atlas Network, Fujimoto J, Scapulatempo-Neto C, Moreira AL, Hwang D, Huang J, Marino M, Korst R, Giaccone G, Gokmen-Polar Y, Badve S, Rajan A, Strobel P, Girard N, Tsao MS, Marx A, Tsao AS and Loehrer PJ. The integrated genomic landscape of thymic epithelial tumors. Cancer Cell 2018; 33: 244-258, e210.

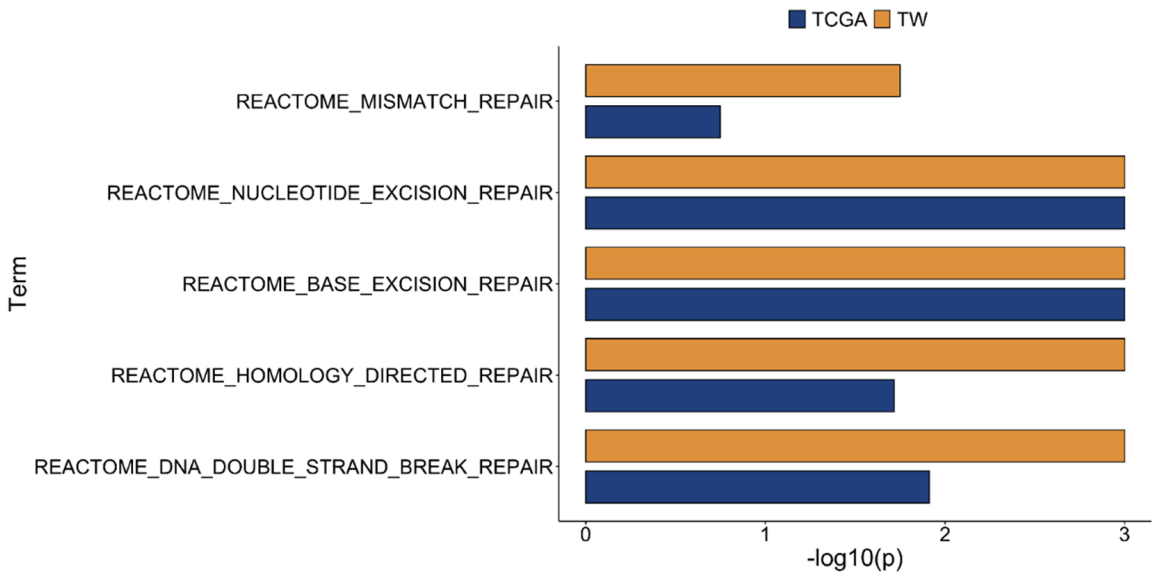
[40] Munoz-Maldonado C, Zimmer Y and Medova M. A comparative analysis of individual RAS mutations in cancer biology. Front Oncol 2019; 9: 1088.

[41] Marx A, Belharazem D, Lee DH, Popovic ZV, Reissfelder C, Schalke B, Scholch S, Strobel P,

Weis CA and Yamada Y. Molecular pathology of thymomas: implications for diagnosis and therapy. Virchows Arch 2021; 478: 101-110.

[42] Bian D, Zhao L, Zhang X, Lv F, Zhu Z, Qiu H and Zhang H. Immunohistochemistry biomarker TP53 expression predicts the survival of thymomas. Gland Surg 2020; 9: 291-299.

[43] Lawrence L, Kunder CA, Fung E, Stehr H and Zehnder J. Performance characteristics of mutational signature analysis in targeted panel sequencing. Arch Pathol Lab Med 2021; 145: 1424-1431.



Supplementary Figure 1. DNA repair pathways of Taiwanese population showed more activity than in TCGA population.