ESMO GOOD SCIENCE BETTER MEDICINE BEST PRACTICE

ESMO OPEN SCIENCE FOR OPTIMAL CANCER CARE

ORIGINAL ARTICLE

Somatic and germline genomic variation in thymic carcinomas

F. Tettamanzi[1], C. De Carlo[2,3], C. Cappadona[1], D. Giuliano[4], V. Rimoldi[1], C. Piombo[3], L. G. Cecchi[1,2], S. Zanella[3], G. Marulli[2,5], E. Voulaz[2,5], G. Veronesi[6,7], S. Marchini[8], L. Di Tommaso[2,3], P. A. Zucali[2,9†] & R. Asselta[1,2†*]

[1]IRCCS Humanitas Research Hospital, Milan; [2]Department of Biomedical Sciences, Humanitas University, Milan; [3]Pathology Unit, IRCCS Humanitas Research Hospital, Milan; [4]Translational Genomic Unit, IRCCS Humanitas Research Hospital, Milan; [5]Division of Thoracic Surgery, IRCCS Humanitas Research Hospital, Milan; [6]School of Medicine, Vita-Salute San Raffaele University, Milan; [7]Department of Thoracic Surgery, IRCCS San Raffaele Scientific Institute, Milan; [8]Laboratory of Cancer Pharmacology, IRCCS Humanitas Research Hospital, Milan; [9]Department of Oncology, IRCCS Humanitas Research Hospital, Milan, Italy

Available online 10 March 2026

Background: Thymic carcinoma (TC) is a rare and aggressive neoplasm of thymic epithelial origin, with no currently available biomarkers to guide treatment or prognosis. In this study, we aimed at characterizing the genomic landscape of a cohort of TCs, at somatic and germline levels.

Materials and methods: Whole-exome sequencing was carried out on 25 TC archival samples and 17 histologically normal adjacent tissues. Genomic characteristics and their association with clinical outcomes were assessed.

Results: High tumor mutational burden (TMB) and microsatellite instability were present in 8% and ∼30% of the study cohort. Somatic mutations in 20 core genes of DNA damage response (DDR) were associated with high TMB ($P = 0.005$) and worse overall survival (OS) [hazard ratio 5.05 (95% confidence interval 1.49-17.20), $P = 0.010$]. Putative pathogenic variants in the core DDR genes were frequently detected in the germline as well (35%). Among somatic alterations, driver genes included *POLE*, *MTOR*, and *ATR* (12%). Notably, ∼30% of patients harbored mutations matching biomarkers of response to currently available therapies such as poly (ADP-ribose) polymerase inhibitors. Somatic copy number profiles were heterogeneous; however, several amplification and deletion events were identified as recurrent across tumors.

Conclusions: TC appears to be a heterogeneous disease, exhibiting varying genomic alteration rates. Several somatic and germline aberrations, some of which have been reported in this article for the first time, warrant further investigation for their hinted prognostic value and clinical implications.

Key words: thymic carcinoma, somatic and germline variants, tumor mutational burden, microsatellite instability, copy number alterations

INTRODUCTION

Malignancies originating from thymic epithelial cells are known as thymic epithelial tumors (TETs). They are a rare and heterogeneous group of tumors, with thymic carcinoma (TC) being one of the rarest forms, affecting 0.2-0.5 new cases per million person-years.[1] Among TETs, TCs display a more aggressive clinical behavior, characterized by a higher risk of relapse and distant dissemination. Patients' prognosis is poor, resulting in a 5-year overall survival (OS) rate of 36%.[2] Risk factors are unknown, and the understanding of tumor etiology is hampered by its low incidence. Surgical resection is the cornerstone treatment for early-stage TCs, whereas for advanced and metastatic tumors chemotherapy or radiochemotherapy is indicated[3] to achieve disease control. Platinum-based chemotherapy is administered as first line, and no standard is defined for subsequent lines. Few novel therapies have been investigated in advanced and metastatic settings[4] in the last decade, and targeted drugs—such as lenvatinib,[5] sunitinib,[6] and everolimus[7]—as well as immune checkpoint inhibitors—such as pembrolizumab[8]—have shown response rates up to 30%, appearing promising yet still requiring validation in large trials.

Tumor genomic profiling entails unevaluable opportunities for the identification of biomarkers, development of targeted therapies, and investigation of tumor pathogenesis. Germline genomic profiling in cancer patients has additionally emerged as a fundamental approach to assess heritable cancer risk[9] and inform about tumor vulnerabilities, as in the case of homologous recombination deficiency upon poly (ADP-ribose) polymerase (PARP) inhibition.[10] Clinically relevant biomarkers in cancer currently embrace somatic

*Correspondence to: Prof. Rosanna Asselta, Department of Biomedical Sciences, Humanitas University, Via Rita Levi Montalcini 4, 20072 Pieve Emanuele, Milan, Italy. Tel: +39-02-82245215
E-mail: rosanna.asselta@hunimed.eu (R. Asselta).

†These authors share last authorship.

and germline genetic aberrations, including single-nucleotide variants (SNVs), changes in repetitive sequences, somatic copy number alterations (SCNAs), as well as more complex chromosomal rearrangements. In TC, recurrent mutations in genes involved in cell cycle regulation (such as *TP53* and *CDKN2A*), known oncogenic pathways (such as *RAS*, *KIT, PI3CA*, and *MTOR*), as well as epigenetic mechanisms (including *BAP1, SETD2, DNMT3*, and *TET2*)[11] have been described. Chromosomal arm-level aberrations are also common, mostly losses at chromosome 16q.[12,13] Analysis of a TET cohort including 11 TCs by The Cancer Genome Atlas identified a TC-like cluster characterized by higher tumor mutational burden (TMB), aberrant expression of MYC/Max, MYB, p53, XBP1-2, and PPARA/RXRA, as well as loss at chromosome 16q.[12] Cases of high TMB as well as high microsatellite instability (MSI) due to somatic and germline mutations in the mismatch repair (MMR) genes have also been reported,[12,14] supporting—along with high programmed death-ligand 1 expression in 70% of tumors[15]—the potential role of immune checkpoint inhibitors in TC treatment. Despite these advancements, TC characterization relies primarily on small-scale studies that often focus on a narrow set of cancer-related genes. Thus, the pathogenesis of TC is far from being fully elucidated and the availability of biomarkers for patients' prognosis and treatment represents an unmet need. In this article, we provide a comprehensive somatic and germline profiling through whole-exome sequencing (WES) with the aim of characterizing the genomic landscape of a cohort of TC patients and assessing possible biomarkers and their association with clinical outcomes.

MATERIALS AND METHODS

Patients and samples

Formalin-fixed paraffin-embedded (FFPE) samples derived from resected TCs ($n = 25$) and histologically normal adjacent tissues ($n = 17$) were retrieved from the Humanitas Clinical and Research Center (Milan, Italy). Patients' demographic and clinical data were collected, including age, sex, dates of diagnosis and surgery, World Health Organization histological classification,[16] TNM (tumor—node—metastasis) staging,[17] and dates of relapse, death, and last observation. Blood samples of unrelated non-cancer patients ($n = 27$) were included to build a normal reference for genomic analyses. The study was conducted in accordance with the Good Clinical Practice guidelines, the ethical principles of the Declaration of Helsinki, and local regulations. Patients signed an informed consent, and all procedures were approved by the appropriate institutional committee (protocol ONC/OSS-09/2021 approved by the Internal Review Board on 23 February 2021).

Control cohort

The control cohort derived from the Atherosclerosis, Thrombosis, and Vascular Biology Italian Study Group[18] composed of 1755 individuals recruited between 1998 and 2001 by 94 Italian hospitals, belonging to the Myocardial Infarction Genetics Consortium[19] (MIGen). Details about samples and WES are reported in Mannucci et al.[18] WES data of these healthy controls are deposited in the National Institutes of Health (NIH) dbGAP repository under accessions phs000279 and phs000806. Given the limited number of cases with available normal tissue ($n = 17$) in our cohort, a subgroup of controls ($n = 102$) was selected for comparison, matching TC patients by sex and age (Supplementary Figure S2D, available at https://doi.org/10.1016/j.esmoop.2026.106305). Matching was carried out using the MatchIt R package[20] to obtain a case : control ratio of 1 : 6 by applying the formula phenotype $\sim$ sex $+$ age.

Whole-exome sequencing

Genomic DNA (gDNA) was extracted from FFPE samples, using the Maxwell® RSC DNA FFPE Kit (Promega Corporation, Madison, WI) on a Maxwell® RSC instrument (Promega Corporation); gDNA extraction from blood samples was carried out using a Microlab STAR Liquid Handling System (Hamilton, Bonaduz, Switzerland) integrated with Chemagen automated DNA extraction system (Chemagen AG, Baesweiler, Germany). Exome capture was carried out with the SureSelectXT HS Target Enrichment Kit (Human All Exon v8) using the Magnis NGS Prep System (Agilent Technologies, Santa Clara, CA) according to the manufacturer's protocol. Libraries were sequenced as 150-base-pair paired-end reads on a NextSeq2000 instrument (Illumina, San Diego, CA).

Alignment and SNV/indel calling

Reads were aligned to the reference genome GRCh38 using BWA-mem[21] v.0.7.17. Base quality was recalibrated according to Genome Analysis ToolKit[22] Best Practices v.4.4.2.0. Somatic variants were called using Mutect2[23] v.2.2 with default parameters. A panel of normals was generated running Mutect2 on data from control blood and tissue normal samples and including variants detected in two or more samples. A consensus call was obtained using Strelka2[24] v.2.9.10 for paired samples and VarDict-java[25] v.1.8.3 for tumor-only samples. Germline variants were called using HaplotypeCaller[26] v4.2.2.0 with default parameters, and a consensus call was obtained using Strelka2 v2.9.10. Applied filters are reported in Supplementary Materials and Methods, available at https://doi.org/10.1016/j.esmoop.2026.106305.

Variant annotation and pathogenicity prediction

Variants were annotated with Ensembl Variant Effect Predictor[27] v.104. Pathogenicity on protein function was predicted using Rare Exome Variant Ensemble Learner[28] (REVEL) released on February 2020. Parameters to define putative pathogenic variants are reported in the Supplementary Materials and Methods, available at https://doi.org/10.1016/j.esmoop.2026.106305. Driver mutations

were predicted using Cancer Genome Interpreter[29] (CGI) released on 17 January 2018.

TMB and MSI status

TMB was defined as the number of somatic, coding, non-synonymous SNVs/indels per megabase (Mb) of the examined genome. Samples were defined as TMB-high (H) if TMB was $\geq$10 mut/Mb. Somatic MSI status was determined using MSIsensor-pro[30] v.1.2.0 with default parameters. MSI score was calculated as a percentage of unstable sites/total sites. Tumors were defined as MSI-H if the MSI score was $\geq$10%, MSI-intermediate (MSI-I) if the MSI score was 3%-10%, and MSI-low (MSI-L) if the MSI score was $<$3%.[31]

SCNA calling

PureCN[32] v.2.4.0 was used to estimate purity and ploidy and infer absolute copy number (CN) states of tumors according to PureCN best practices. For allele-specific estimation, germline heterozygous single-nucleotide polymorphisms were called using Mutect2 v.2.2 in tumors and confirmed through Haplotypecaller v.4.2.2.0 in matched-normal samples as applicable. Flagged purity and ploidy solutions were manually curated. Additional details are reported in Supplementary Materials and Methods, available at https://doi.org/10.1016/j.esmoop.2026.106305.

SCNA analyses

CN profiles were summarized as: (i) number of segments constituting CN profiles; (ii) SCNA burden (i.e. fraction of genome-bearing CN events); (iii) number of segments in loss of heterozygosity (LOH); (iv) LOH burden (i.e. fraction of genome in LOH). Recurrent SCNAs were identified using Genomic Identification of Significant Targets in Cancer 2[33] (GISTIC2) v.2.0.23. Additional details are reported in Supplementary Materials and Methods, available at https://doi.org/10.1016/j.esmoop.2026.106305.

Mutational and CN signatures

Mutational and CN signatures were extracted from tumor mutational and CN profiles using MutationalPattern[34] v.3.4 and Sigprofiler toolkit[35] v.3.3, respectively. For both mutational and CN signatures, a refitting approach based on COSMIC mutational signatures[36] v.3.3 was used. Exposure was estimated for each sample and reconstruction performance was evaluated using cosine similarity between original and fitted profiles.

Gene lists and enrichment analyses

Genomic alterations in genes involved in DNA damage response (DDR) and other cancer-related processes were assessed using: (i) a core list of 20 DDR genes associated with autosomal dominant cancer-predisposition syndromes involving maintenance of DNA integrity[9]; (ii) a more comprehensive non-redundant DDR gene list assembled from relevant Molecular Signature Database[37] v.2023.1 gene sets; (iii) the COSMIC cancer gene census list[36] (CGC) v.97. Over-representation analysis was carried out using Metascape[38] v.3.5 with default parameters, a background genome-wide gene list, and terms selected from Gene Ontology. Gene lists are provided in Supplementary Table S1, available at https://doi.org/10.1016/j.esmoop.2026.106305.

Actionability analysis

Alterations therapeutically actionable or predictive of treatment response were identified through CGI[29] released on 17 January 2018. Analysis was limited to alterations perfectly matching biomarkers of drug sensitivity and to evidence levels A-C: A—Food and Drug Administration (FDA) and National Comprehensive Cancer Network (NCCN) guidelines; B—late-phase trials; C—early-phase trials and case reports. CGI annotations were crossed with COSMIC actionability list v.17 and verified with the FDA website and source drug labels for marketed compounds.

Statistical analyses

Statistical analyses were carried out using R v.4.2.3 (https://www.R-project.org/https://www.R-project.org/). Continuous variables were described as median with interquartile range (IQR) or mean with standard deviation (SD), whereas categorical variables were described as frequency and percentage of units/level. Association between continuous or ordinal variables was assessed using Spearman's correlation ρ. Medians of unrelated groups were compared using the Wilcoxon (W) rank-sum test,[39] while relationship between categorical variables was assessed using the chi-square test. Tests were two-sided at nominal level $\alpha = 0.05$. False discovery rate was applied for multiple testing correction in high-dimensional analyses as indicated, and q values <0.05 were considered statistically significant. In survival analysis, OS time was defined as the interval between the date of surgery or histological diagnosis, whichever occurred first, and date of death from any cause. For patients remaining alive, time was censored at the date of last observation. Survival curves and medians with 95% confidence interval (CI) were estimated using the Kaplan—Meier (KM) method[40] and compared by log-rank test.[41] Median follow-up time was estimated using the reverse KM method.[42] Cox regression[43] was applied to assess associations with OS in terms of hazard ratios (HRs) with 95% CI. Clinical parameters were evaluated as possible confounders assessing association with survival through log-rank test and simple Cox regression models. In multiple regression, only associated clinical parameters (i.e. complete resection) were included in the models as covariates. Proportional-hazards assumption was assessed testing the independence of scaled Schoenfeld residuals of time.

RESULTS

Patients' characteristics

The study cohort included 25 patients diagnosed with TC, who underwent surgery between January 2003 and December 2022. Baseline clinical information is provided in

Table 1. Demographic and clinical characteristics of the study cohort

Characteristics	Patients (n = 25)
Age at diagnosis, median years (IQR)	60 (51-70)
Sex, % (*n*)	
Female	36 (9)
Male	64 (16)
WHO histological type, % (*n*)	
Squamous thymic carcinoma	60 (15)
Basaloid thymic carcinoma	20 (5)
Lymphoepithelial thymic carcinoma	4 (1)
Mucoepidermoid thymic carcinoma	4 (1)
Thymic carcinoma, NOS	8 (2)
Undifferentiated thymic carcinoma	4 (1)
Stage at diagnosis, % (*n*)	
I	16 (4)
II	8 (2)
III	40 (10)
IV	36 (9)
Surgery, % (*n*)	
Yes	100 (25)
Residual disease after surgery, % (*n*)	
Yes	20 (5)
No	80 (20)
Neoadjuvant treatment, % (*n*)	
Yes	40 (10)
No	60 (15)
Adjuvant treatment, % (*n*)	
Yes	64 (16)
No	36 (9)

IQR, interquartile range; NOS, not otherwise specified; WHO, World Health Organization.

Table 1. Patients had a median age at diagnosis of 60 years (IQR 51-70 years) and were predominantly males (64%). The most common histological type was squamous carcinoma (60%), followed by basaloid carcinoma (20%). At the time of diagnosis, patients presented mainly stage III and IV tumors (76%). The median duration of follow-up was 5.01 years (95% CI 4.13- not available) since the time of surgery, while the median OS was 4.57 years (95% CI 3.55-not available).

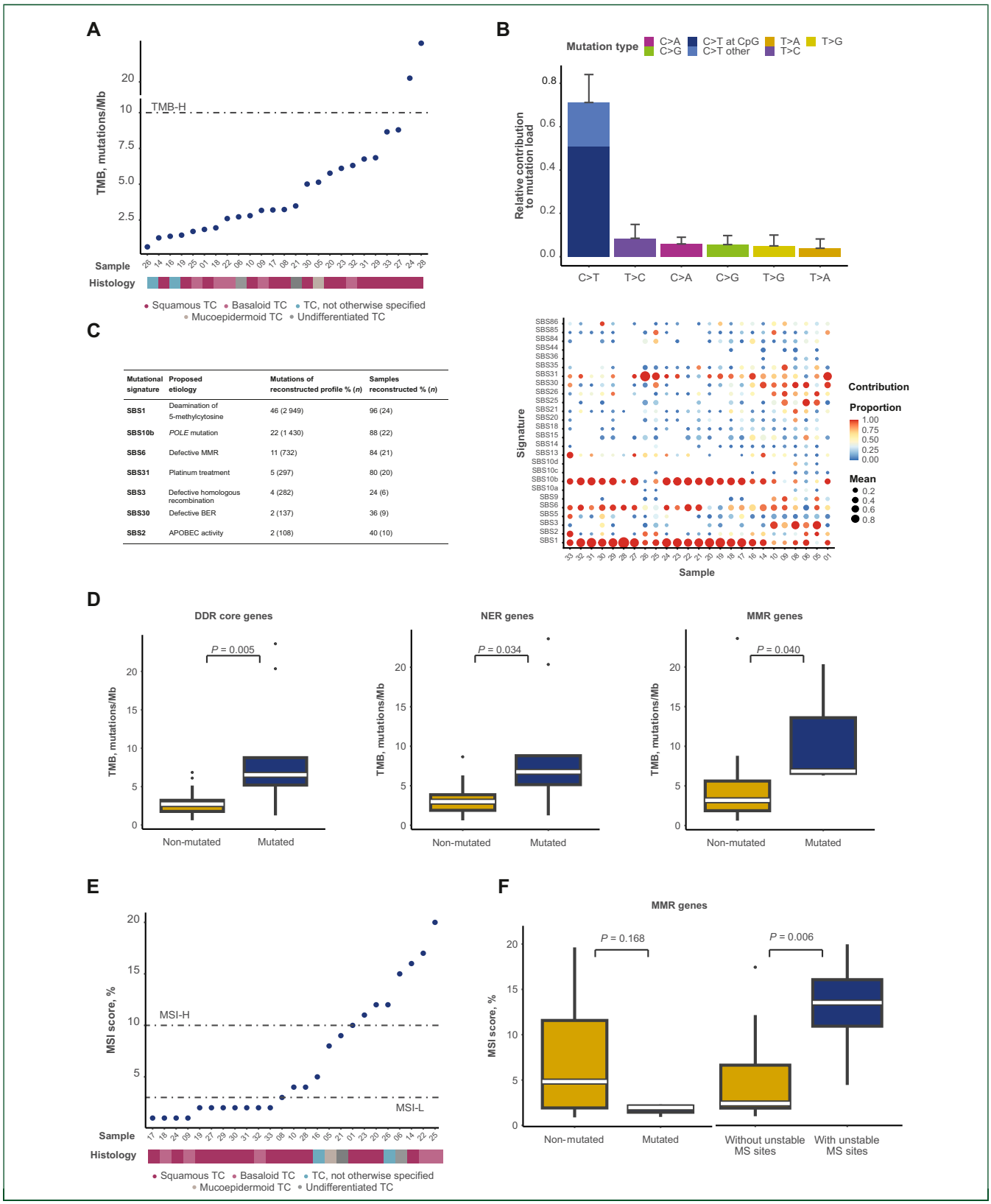
Somatic mutation profile

We carried out WES on 25 TCs and 17 matched-normal adjacent tissues. DNA was sequenced to a median coverage of 176X (IQR 97X-200X) and altogether 6471 high-quality variants were detected (Supplementary Figure S1, available at https://doi.org/10.1016/j.esmoop.2026.106305). The genomic landscape of analyzed TCs, including the biomarkers TMB and MSI, the distribution of base substitutions, and mutational signatures, is shown in Figure 1. C>T transitions prevailed in TC mutational spectra (77% of mutations), especially at CpG islands (62% of mutations). Other types of base mutations were far less frequent ($\leq$8% for each class). Tumors displayed highly heterogeneous mutational loads with few tumors showing very low or very high mutation frequencies (TMB <1 mut/Mb: 4%; TMB >20 mut/Mb: 8%). The median TMB was 5.39 mut/Mb (IQR 1.95-6.31 mut/Mb), whereas the prevalence of TMB-H tumors was 8% (Figure 1A).

To investigate the biological processes underlying mutation onset, we assessed the presence of mutational patterns within tumor profiles. Referring to known mutational signatures reported in the COSMIC catalogue, 23 signatures were detected to different extents across tumors (reconstruction accuracy: median cosine similarity 0.93, IQR 0.88-0.95) (Figure 1C). Among the biological processes related to the identified signatures, deamination of 5-methylcytosine contributed the most (46% of mutations) and was ubiquitous (96% of tumors). Six other signatures contributed with >100 mutations each to the overall mutational load and were frequently detected (>30% of samples, except for defective HR repair), related to *POLE* mutation, defects in several DNA repair systems [MMR, HR, and base excision repair (BER)], chemotherapy treatment, and APOBEC deaminase activity. A positive correlation between TMB and signatures related to MMR deficiency (SBS6 $\rho = 0.689$, $P < 0.001$; MMR signatures combined $\rho = 0.674$, $P < 0.001$) as well as to replicative DNA polymerases (*POLE* SBS10b $\rho = 0.707$, $P < 0.001$; replicative *POL* signatures combined $\rho = 0.795$, $P < 0.001$) was observed.

Mutation load in DDR and cancer-related genes

Detection of high mutational loads and specific signatures led us to investigate the presence of somatic mutations affecting DNA damage signaling and repair systems. Half of the tumors harbored mutations in the 20 core DDR genes and ~20% carried at least one mutation predicted to have a highly deleterious impact on the protein function. Tumors bearing alterations in DDR genes displayed higher TMB compared with tumors without genetic variants (W test $P = 0.005$; $\rho = 0.577$, $P = 0.003$) (Figure 1D), and significant association between increased TMB and the presence of MMR (W test $P = 0.040$; $\rho = 0.427$, $P = 0.033$) and NER alterations (W test $P = 0.034$; NER genes $\rho = 0.439$, $P = 0.028$) was observed; no difference was detected with BER alterations (W test $P = 0.785$, Supplementary Figure S1A, B, and E, available at https://doi.org/10.1016/j.esmoop.2026.106305). A positive correlation was also observed between TMB and alterations in genes involved in DNA replication ($\rho = 0.626$, $P = 0.001$). Non-synonymous mutations were detected in all DNA repair pathways considered, including NER (44%), BER (40%), MMR, and HR (20%), as well as in DDR signaling pathways (32%). MMR-related signatures were supported by mutations in MMR genes (SBS6 $\rho = 0.513$, $P = 0.009$; MMR signatures combined $\rho = 0.495$, $P = 0.012$), while signature of *POLE* SBS10b correlated positively with alterations in genes involved in DNA replication machinery ($\rho = 0.425$, $P = 0.034$; replicative *POL* signatures combined $\rho = 0.570$, $P = 0.003$). We further investigated genes predicted as drivers according to the CGI (5% of all mutations), identifying *POLE*, *MTOR*, *ATR*, *SVEP1*, *NSD2*, *ERCC4*, and *DKC1* (12%) as most frequently mutated genes. Along with DNA damage repair, gene ontology analysis pointed to the presence of aberrations in several oncogenic processes including chromatin organization, cell cycle progression, p53 regulation, and cell apoptosis. The list of most frequently altered genes involved in DDR and other cancer-related pathways is displayed in Figure 2.



Mutational signature	Proposed etiology	Mutations of reconstructed profile % (*n*)	Samples reconstructed % (*n*)
SBS1	Deamination of 5-methylcytosine	46 (2 949)	96 (24)
SBS10b	*POLE* mutation	22 (1 430)	88 (22)
SBS6	Defective MMR	11 (732)	84 (21)
SBS31	Platinum treatment	5 (297)	80 (20)
SBS3	Defective homologous recombination	4 (282)	24 (6)
SBS30	Defective BER	2 (137)	36 (9)
SBS2	APOBEC activity	2 (108)	40 (10)

Figure 1. Somatic mutational profile and microsatellite instability in thymic carcinomas. TMB (A); mutational spectrum (B); active COSMIC SNV signatures (C); association between TMB and somatic mutations in 20 core DDR, NER, and MMR genes (D); MSI score (E); association between MSI score and somatic mutations and unstable MS sites in MMR genes (F). In C, only signatures with absolute fitted mutational load >100 variants and with >1% contribution are reported in the table. In D and F, difference between groups was assessed using the Wilcoxon test. $n = 25$.

APOBEC, apolipoprotein B mRNA editing enzyme catalytic polypeptide-like; BER, base excision repair; DDR, DNA damage response; H, high; L, low; MMR, mismatch repair; NER, nucleotide excision repair; *POLE*, polymerase epsilon; SBS, single base substitution signature; TC, thymic carcinoma; TMB, tumor mutational burden.

https://doi.org/10.1016/j.esmoop.2026.106305

MSI profile

We next investigated the phenotype of MSI, where genetic alterations are related to the variable length of short repetitive DNA sequences namely microsatellites (MS). The degree of MSI, in terms of percentage of unstable/total number of MS considered, was highly heterogeneous. The median MSI score was 4.1%, (IQR 1.9%-11.1%), and ∼30% of tumors were classified as MSI-H (MSI score >10%) (Figure 1E). An inverse relationship with TMB was observed ($\rho = -0.440$, $P = 0.028$), and no significant difference in MSI scores was observed between tumors mutated in MMR genes with respect to wild-type tumors (W test $P = 0.168$; $\rho = -0.29$, $P = 0.160$) (Figure 1F). We further analyzed the somatic unstable MS identified and the affected genes, since the presence of these sites could disrupt gene function when falling in the coding sequence. Notably, unstable MS were identified in exons of several DDR and other cancer-related genes (Supplementary Figure S1F, available at https://doi.org/10.1016/j.esmoop.2026.106305), and a significant difference in MSI was observed between tumors with and without unstable MS in MMR genes (W test $P = 0.006$; $\rho = 0.571$, $P = 0.003$) (Figure 1F).

SCNA profile

We extracted tumor CN profiles to determine the presence and extent of changes in DNA copies. Analysis of genomic SCNAs in TCs, including CN signatures as well as regions of recurrent genomic imbalance, is shown in Figure 3. Tumor CN profiles were first summarized in terms of SCNA and LOH burden indicating the proportion of genomic regions affected by SCNAs and LOH, respectively (Figure 3A and B). Median SCNA burden was 0.23 (IQR 0.16-0.29), while median LOH burden was 0.11 (IQR 0.05-0.14). Few tumors were characterized by genomes without CN and LOH events (SCNA and LOH burden ∼0: 12%), and few displayed a higher degree of alterations (SCNA and LOH burden >0.40: 4%). Advanced tumors (stages III-IV) showed a tendency to accumulate SCNA burden ($P = 0.104$) (Figure 3C) compared with early-stage tumors (stages I-II). A similar trend was observed when stage IV tumors were compared with all other tumors ($P = 0.100$, data not shown). LOH burden was slightly higher in advanced tumors compared with early-stage tumors, but not to a statistically significant extent ($P = 0.355$) (Figure 3D).

To investigate the processes underlying the generation of CN imbalances, we assessed the presence of CN patterns within tumor profiles. Referring to known CN signatures from the COSMIC catalogue, nine signatures were detected to different degrees in tumors (reconstruction accuracy: median cosine similarity of 0.93, IQR 0.91-0.96) (Figure 3E). Samples were characterized mostly by signature of diploid state (CN1: 88%) and, to a lesser extent, by signature of tetraploid state (CN2: 32%), pointing to likely genome duplication events. Patterns of CN alterations were ascribable to focal LOH and chromosome instability, homologous recombination deficiency, and chromosomal LOH in a minority of tumors (CN9: 12%, CN17: 8%, and CN13: 4%).

Among observed SCNAs, several statistically significant recurrent events were identified, which refer to ∼60% of autosomal chromosomes. Regions of recurrent amplification and deletion are displayed in Figure 3G. Assessment of genetic loci revealed the presence of two known oncogenes in amplified regions, *MUC4* and *COL11A1*, both involved in the cell adhesion and extracellular matrix remodeling and signaling. In deleted regions, several tumor suppressors were identified including genes involved in the regulation of cell cycle (*CDKN2A*, *STK11*, and *ZFHX3*), DNA damage repair and response (*FANCA* and *RFWD3*), chromatin organization and transcriptional regulation (*CTCF*, *CBFA2T3*, *CBFB*, *TCF3*, and *ZFHX3*), as well as cell−cell adhesion (*CDH1* and *CDH11*).

Germline mutation profile

To assess the presence of germline variants related to tumor susceptibility, we carried out germline variant calling on a subset of 17 histologically normal matched thymic tissues. Overall, a total of 331 799 high-quality variants were detected, with a median load of 19 232 variant sites/sample (IQR 14 075-21 129) (Supplementary Figure S2A-D, available at https://doi.org/10.1016/j.esmoop.2026.106305). A detailed representation of TC frequently altered genes involved in DDR and other cancer-related pathways is displayed in Figure 4. Germline variants affecting the 20 DDR core genes were identified in all analyzed samples, for a total of 33 detected non-synonymous putative pathogenic variants in 47% of patients. Notably, ∼35% of patients harbored at least one germline variant predicted to have a highly deleterious impact on protein function. Similar to somatic alterations, non-synonymous variants were detected in all considered DNA repair pathways including NER (71% of tumors), HR (65%), MMR, and BER (59%), as well as in DDR signaling pathways (76%). Thereof, only two variants in *POLD1* and *FAS* genes were already reported in ClinVar[44] with a clinical annotation of uncertain significance. To provide a basis for evaluation, we analyzed germline non-synonymous SNVs in a control cohort ($n = 102$) from the Italian general population, matched to TC patients by age and sex (Supplementary Figure S2E, available at https://doi.org/10.1016/j.esmoop.2026.106305). In the control group, variants in the DDR core genes were detected at lower prevalence (non-synonymous set: 25% versus 47%, $P = 0.075$; deleterious set: 8% versus 35%, $P = 0.006$) and with a lower fraction per individual [mean (SD) non-synonymous set: 1.12 (0.44) versus 4.12 (3.80), $P < 0.001$; deleterious set: 0.36 (0.57) versus 1.88 (1.81), $P = 0.008$]. The frequency of co-mutating genes per individual was also lower [mean (SD) non-synonymous set: 0.20 (0.71) versus 2.38 (2.00), $P < 0.001$; deleterious set: 0.08 (0.4) versus 0.75 (1.39), $P = 0.070$]. Similar differences were even more apparent when considering deleterious variants in the extended DDR gene set [prevalence: 73% versus 100%, $P = 0.024$; mean (SD) variants per individual: 1.08 (0.91) versus 5.94 (8.21), $P < 0.001$; mean (SD) co-mutating genes: 0.59

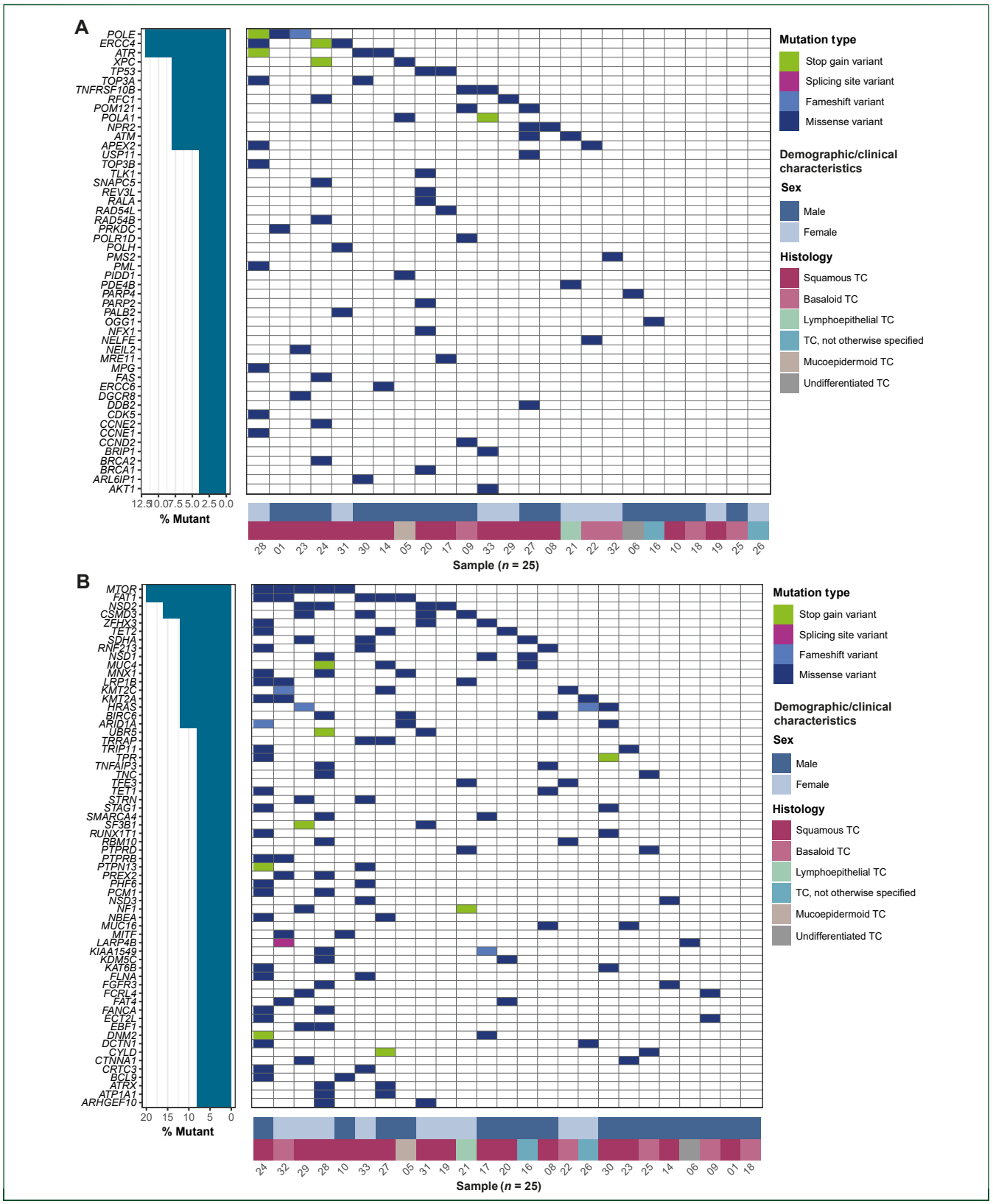
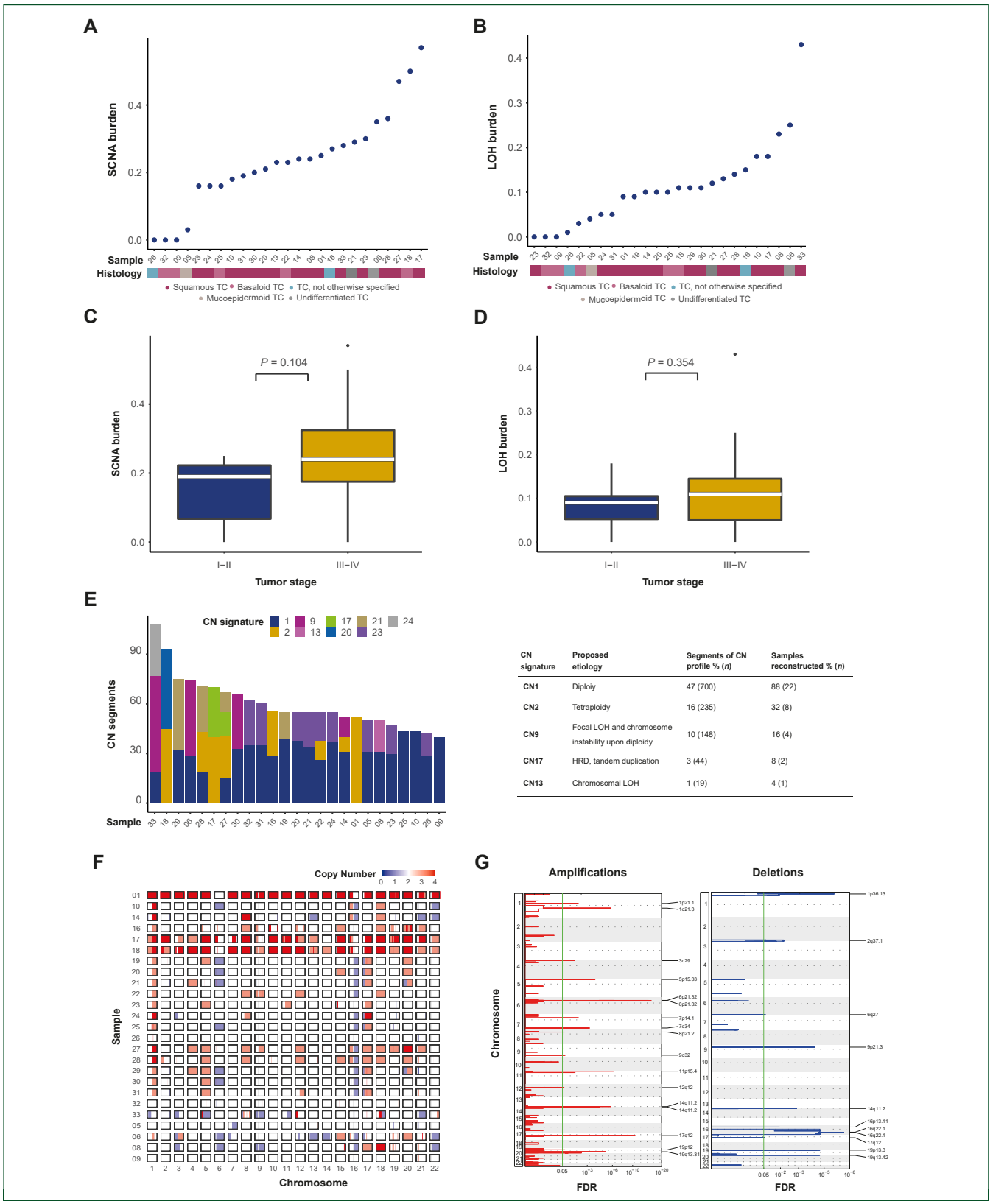


Figure 2. Most frequent somatic gene alterations in thymic carcinomas. Waterfall plot of somatic single-nucleotide variants affecting DDR genes (A) and non-overlapping COSMIC Cancer Gene Census genes (B). Only non-synonymous variants not previously annotated in ClinVar as benign/likely benign or common in the reference population according to gnomAD were considered. In B, mutated genes with frequency ≥8% in the study cohort are displayed. $n = 25$.
F, female; M, male; TC, thymic carcinoma.



CN signature	Proposed etiology	Segments of CN profile % (*n*)	Samples reconstructed % (*n*)
CN1	Diploiy	47 (700)	88 (22)
CN2	Tetraploidy	16 (235)	32 (8)
CN9	Focal LOH and chromosome instability upon diploidy	10 (148)	16 (4)
CN17	HRD, tandem duplication	3 (44)	8 (2)
CN13	Chromosomal LOH	1 (19)	4 (1)

Figure 3. Somatic copy number alterations in thymic carcinomas. SCNA burden (A); LOH burden (B); association between tumor stage and SCNA (C) and LOH burden (D); active COSMIC CN signatures (E); heatmap of SCNA per sample (F); GISTIC2 significantly amplified and deleted genomic regions with FDR *q* value <0.05 (G). In C and D, difference in scores between groups was assessed using the Wilcoxon test. In F and G, red and blue colors indicate CN gains and losses, respectively. $n = 25$.

CN, copy number; FDR, false discovery rate; HRD, homologous recombination deficiency; LOH, loss of heterozygosity; SCNA, somatic copy number alteration; TC, thymic carcinoma.

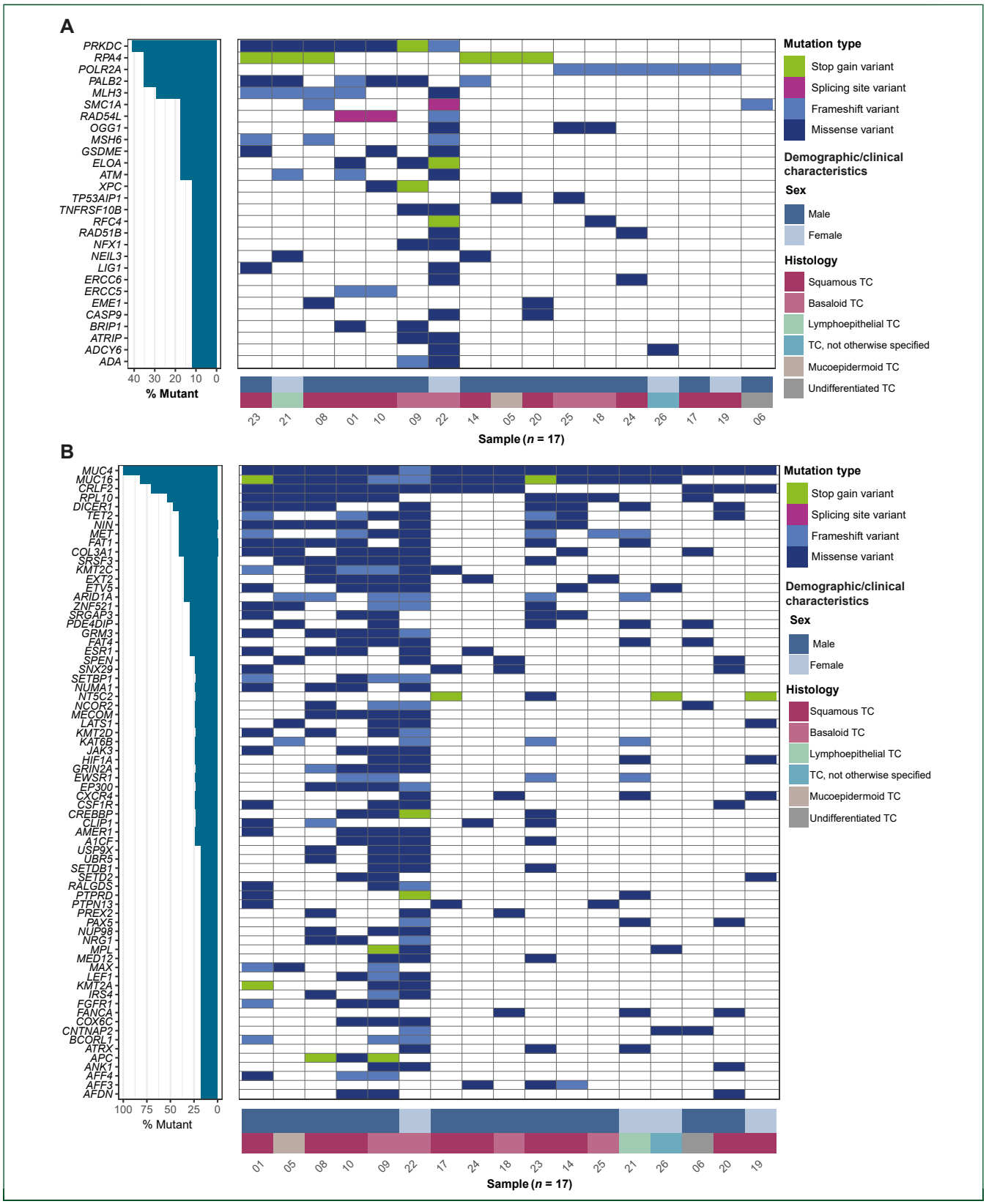


Figure 4. Most frequent germline gene alterations in histologically normal matched thymic tissues. Waterfall plot of germline single-nucleotide variants affecting DDR genes (A) and non-overlapping COSMIC Cancer Gene Census genes (B). Only non-synonymous variants not previously annotated as benign/likely benign or common in the reference population according to gnomAD were considered. Mutated genes with frequency ≥10% (A) or ≥15% (B) in the study cohort are displayed. $n = 17$.

F, female; M, male; TC, thymic carcinoma.

(1.05) versus 4.06 (6.4), $P < 0.001$] (Supplementary Figure S2F and G, available at https://doi.org/10.1016/j.esmoop.2026.106305). Together with variation in DDR pathways—especially related to double strand break (DSB) repair, gene ontology analysis evidenced in the TC cohort alterations in oncogenic processes related to chromatin and chromosome organization, cell cycle progression, and p53 transcriptional regulation. Among the top putative drivers identified through CGI in up to 18% of samples, genes were involved in cell functions of epigenetic and transcriptional regulation (*TET2*, *KMT2C*, *ARID5B*, and *CNOT1*), cell cycle regulation (*CDC27* and *HERC2*), as well as cell adhesion and polarity (*CDH8*, *MAGI2*, *MACF1*, and *ANK3*).

Genomic biomarkers of drug response

We investigated repurposing opportunities in terms of off-label therapies for our cohort of TCs, analyzing the presence of genomic biomarkers of sensitivity to currently available treatments (Figure 5A and B). Forty-six unique somatic mutations matching known actionable biomarkers were identified. Considering only those conferring drug sensitivity, ∼76% of patients carried a mutation having a clinical utility supported by evidence derived from clinical guidelines (A), late-phase trials (B), and early-phase trials or case reports (C). Notably, 36% of patients harbored a potential predictive biomarker targeted by marketed drugs used in clinical practice for other indications (level A evidence, Supplementary Table S2, available at https://doi.org/10.1016/j.esmoop.2026.106305). These belonged to two main drug classes, PARP and phosphoinositide 3-kinase (PI3K) inhibitors, targeting mutations in *BRCA2*, *ATM*, *ATR*, and *PIK3CA* genes. Additional 57 unique variants matched known actionable biomarkers when considering germline alterations. About 59% of patients carried a variant conferring drug sensitivity supported by level A-C evidence, whereas 41% harbored a variant targeted by drugs available in the clinic (level A evidence, Supplementary Table S3, available at https://doi.org/10.1016/j.esmoop.2026.106305). Also, in this case, drugs belonged to two main classes, PARP and PI3K inhibitors, targeting mutations in *ATM*, *PALB2*, *CHEK2*, and *PIK3CA* genes among others.

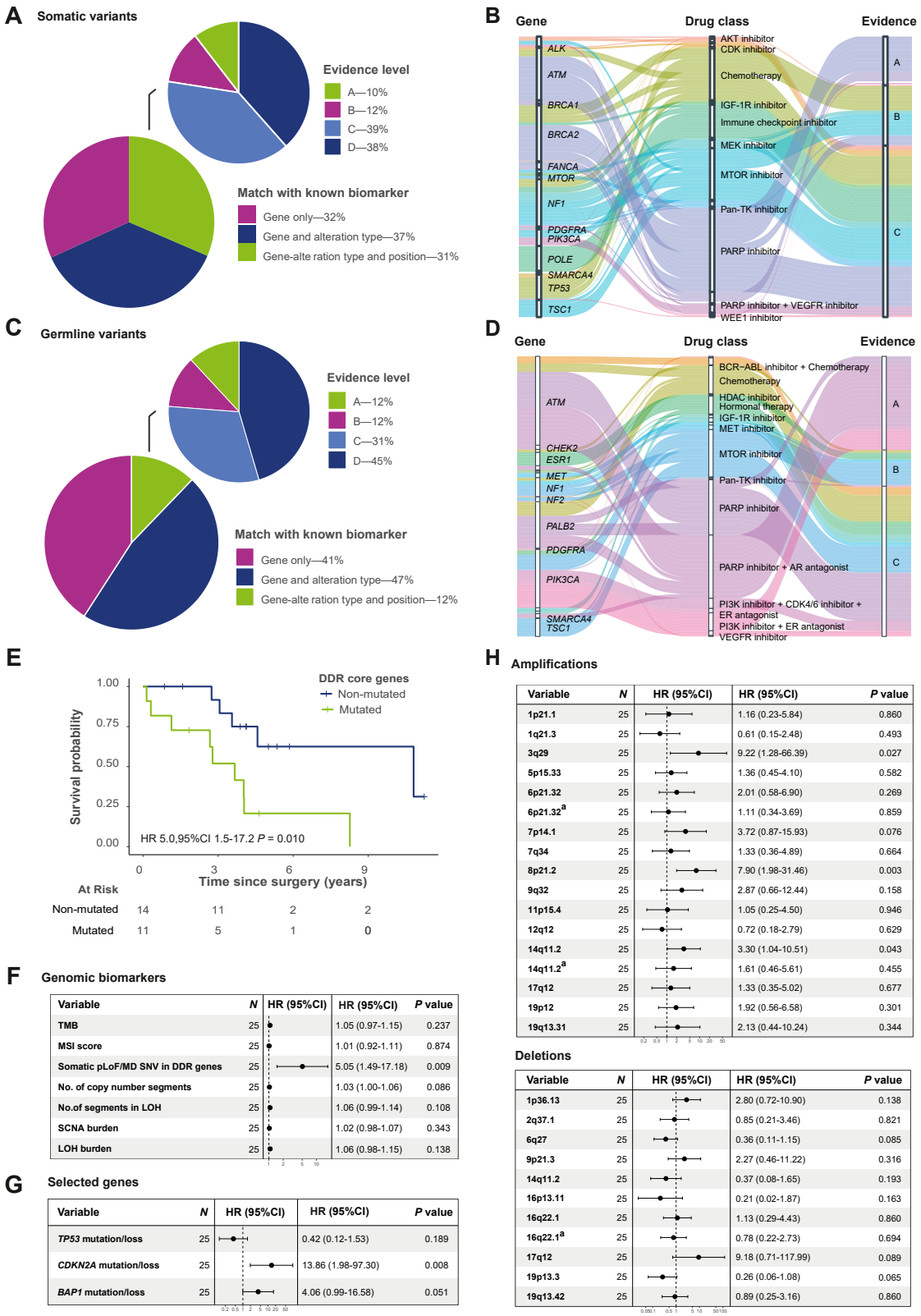
Prognostic value of genomic biomarkers in TC

Lastly, we assessed the prognostic value of the analyzed genomic biomarkers in exploratory survival analyses (Supplementary Table S4, available at https://doi.org/10.1016/j.esmoop.2026.106305). High degree of TMB and MSI did not show a prognostic effect (HR 1.05, 95% CI 0.97-1.15, $P = 0.237$; HR 1.01, 95% CI 0.92-1.11, $P = 0.874$), while the presence of somatic putative pathogenic variants in the 20 DDR core genes appeared to be detrimental (HR 5.05, 95% CI 1.49-17.20, $P = 0.010$) as displayed in the KM curves (Figure 5E and F). No association with survival was detected for *TP53*-mutated or -deleted tumors, whereas a negative effect was shown for *CDKN2A* mutations or deletions (HR 13.9, 95% CI 1.98-97.3, $P = 0.008$) and for *BAP1* mutations or deletions (HR 4.1, 95% CI 0.99-16.6, $P = 0.051$) (Figure 5G). As for SCNAs, a negative trend of association was recorded for the number of segments of CN profiles and in LOH (HR 1.03, 95% CI 0.996-1.06, $P = 0.086$; HR 1.06, 95% CI 0.99-1.14, $P = 0.108$). Effect of recurrent CN events at specific loci was detected, with amplification at chromosomes 3q29, 8p21.2, and 14q11.2 negatively associated with survival (HR 9.2, 95% CI 1.28-66.4, $P = 0.027$; HR 7.9, 95% CI 1.98-31.5, $P = 0.003$; HR 3.3, 95% CI 1.04-10.5, $P = 0.043$) (Figure 5H).

DISCUSSION

In this study, we provided an extensive genomic characterization of a cohort of TC patients, appraising both somatic and germline variation. Our genomic analysis hinted at a profound heterogeneity within TCs, as for the burden of point mutations, unstable MS, as well as SCNAs. Our data confirmed the presence of subgroups of patients displaying high TMB and high MSI. Although the median mutation load was higher than those previously reported[12,45,46] (5.5 versus ∼1.0-3.8 mut/Mb), the portion of TMB-H tumors is in line with the 6%-7% prevalence described in two large consecutive cohorts.[45,47] Tumors with exceptionally high TMB (>20 and >100 mut/Mb) were reported in association with pathogenic mutations in *MLH1* and *MSH6* genes.[12,48] In our cohort, one-fifth of tumors harbored non-synonymous mutations in genes involved in the MMR, and unexpectedly even a higher fraction of patients displayed non-synonymous germline variants, which may be relevant for TC susceptibility. While a few studies on TC have reported MSI analysis,[45,47] recent case reports described TC patients harboring an MSI-H tumor due to germline mutations in the *MLH1* gene and a family history suggestive of Lynch syndrome.[14] In our cohort, biomarkers of MSI and TMB did not correlate positively, suggesting the presence of non-overlapping causes of the two phenotypes. Although this observation deviates from the typical pattern, the presence of 'discordant phenotypes' has already been described in large-scale studies,[49,50] where 30%-89% of MSI-H cases were classified as TMB-low. TMB and MSI score did not exhibit an association with disease prognosis, which might be linked both to tumor-specific characteristics[51] and to the power of our analyses. The presence of TMB-H and MSI-H patients however warrants further assessment of the predictive value of these two biomarkers upon immunotherapy, given that a few phase II trials showed clinical activity[8,52] of immune checkpoint inhibitors in TETs, and pembrolizumab has been recently included in the NCCN guidelines as a recommended option from second-line treatments in TCs.[53] In addition, Möhrmann et al. provided molecular data supporting their application, revealing the presence of two immunologically defined entities in TCs, characterized by high or low tumor-infiltrating lymphocytes and displaying significantly different prognoses.[54]

Somatic mutations were detected in genes involved in all analyzed DDR pathways, and alterations in the core DDR



F Genomic biomarkers

Variable	N	HR (95%CI)	HR (95%CI)	P value
TMB	25		1.05 (0.97-1.15)	0.237
MSI score	25		1.01 (0.92-1.11)	0.874
Somatic pLoF/MD SNV in DDR genes	25		5.05 (1.49-17.18)	0.009
No. of copy number segments	25		1.03 (1.00-1.06)	0.086
No.of segments in LOH	25		1.06 (0.99-1.14)	0.108
SCNA burden	25		1.02 (0.98-1.07)	0.343
LOH burden	25		1.06 (0.98-1.15)	0.138

G Selected genes

Variable	N	HR (95%CI)	HR (95%CI)	P value
TP53 mutation/loss	25		0.42 (0.12-1.53)	0.189
CDKN2A mutation/loss	25		13.86 (1.98-97.30)	0.008
BAP1 mutation/loss	25		4.06 (0.99-16.58)	0.051

H Amplifications

Variable	N	HR (95%CI)	HR (95%CI)	P value
1p21.1	25		1.16 (0.23-5.84)	0.860
1q21.3	25		0.61 (0.15-2.48)	0.493
3q29	25		9.22 (1.28-66.39)	0.027
5p15.33	25		1.36 (0.45-4.10)	0.582
6p21.32	25		2.01 (0.58-6.90)	0.269
6p21.32[a]	25		1.11 (0.34-3.69)	0.859
7p14.1	25		3.72 (0.87-15.93)	0.076
7q34	25		1.33 (0.36-4.89)	0.664
8p21.2	25		7.90 (1.98-31.46)	0.003
9q32	25		2.87 (0.66-12.44)	0.158
11p15.4	25		1.05 (0.25-4.50)	0.946
12q12	25		0.72 (0.18-2.79)	0.629
14q11.2	25		3.30 (1.04-10.51)	0.043
14q11.2[a]	25		1.61 (0.46-5.61)	0.455
17q12	25		1.33 (0.35-5.02)	0.677
19p12	25		1.92 (0.56-6.58)	0.301
19q13.31	25		2.13 (0.44-10.24)	0.344

Deletions

Variable	N	HR (95%CI)	HR (95%CI)	P value
1p36.13	25		2.80 (0.72-10.90)	0.138
2q37.1	25		0.85 (0.21-3.46)	0.821
6q27	25		0.36 (0.11-1.15)	0.085
9p21.3	25		2.27 (0.46-11.22)	0.316
14q11.2	25		0.37 (0.08-1.65)	0.193
16p13.11	25		0.21 (0.02-1.87)	0.163
16q22.1	25		1.13 (0.29-4.43)	0.860
16q22.1[a]	25		0.78 (0.22-2.73)	0.694
17q12	25		9.18 (0.71-117.99)	0.089
19p13.3	25		0.26 (0.06-1.08)	0.065
19q13.42	25		0.89 (0.25-3.16)	0.860

Figure 5. Drug repurposing opportunities and prognostic value of analyzed genomic biomarkers. Somatic (A and B) and germline (C and D) SNVs matching known biomarkers of drug response; KM overall survival curves of patients stratified for the presence of deleterious SNVs in DDR core genes (E); forest plot of HR for overall survival according to genomic alteration markers (F), presence of alterations (deleterious SNVs or LOH events) in selected genes (G), and recurrent SCNAs (H). HR estimates from multiple Cox models are adjusted for the presence of residual disease after surgery ($n = 25$). In A-C, frequency of the type of alteration (compatibility with known biomarkers) and of level of support from clinical and scientific data for biomarkers with perfect match are displayed. In B-D, actionable targets and respective drug classes are displayed for each evidence level. Thickness of bar stacks and flows represents frequency, with only unique alterations considered.

CI, confidence interval; del, deleterious; HR, hazard ratio; KM, Kaplan—Meier; LOH, loss of heterozygosity; MSI, microsatellite instability; SCNA, somatic copy number alteration; som, somatic; SNV, single-nucleotide variant; TMB, tumor mutational burden.

[a]In G, duplicated cytobands refer to different peak coordinates.

https://doi.org/10.1016/j.esmoop.2026.106305

genes were associated with high TMB and poor OS. Most frequently mutated genes were *POLE*, *ATR*, and *ERCC4*, which were identified as putative tumor drivers. Among non-DDR genes, several genes were already reported to be altered in TC, including players involved in signaling pathways and growth control (*MTOR* and *HRAS*), epigenetic regulation (*TET2*, *KMT2A*, and *KMT2C*), and cell adhesion (*FAT1, LRP1B*, and *MUC4*).

Genes implicated in DDR harbored variants with high prevalence in the germline as well, with one-third of patients carrying at least one germline putative pathogenic variant in the core genes. Frequent alterations were observed in genes involved in DSB repair (*ATM*, *PALB2, RAD54L*, and *PRKDC*), MMR (*MLH3* and *MSH6*), and BER (*OGG1*), among others. In a control cohort representing the Italian general population, germline variants in the same genes were found at lower prevalence, and in individuals carrying variants the overall burden of mutations was lower both in terms of fraction per individual and of frequency of co-mutating genes, especially when considering the most pathogenic variants. To our knowledge, our data on germline variants represent one of the few reports available in TC together with the results from Möhrmann et al., which were published during the revision of this manuscript. In this study, an 18% prevalence of likely pathogenic variants was reported for a cohort of 76 TET patients including 45 TC patients.[54] Upon analysis of 141 cancer predisposing genes, TCs displayed likely pathogenic variants in *BAP1, BRCA1, DDX41, TP53, FANCD2*, and *MUTYH* genes. Across cancer patients, the presence of germline pathogenic or likely pathogenic variants appears to be ubiquitous. Data regarding prevalence span from ∼7%-8%[55,56] up to 30%[57] of cases, with differences in frequencies likely ascribable not only to single tumor types, but also to gene selection as well as filtering criteria. Among somatic and germline alterations, variants sensitizing to currently available drugs were identified in 36% and 41% of patients, respectively. For these alterations, off-label therapies included mainly PARP inhibitors such as olaparib, niraparib, and rucaparib, due to mutations in multiple genes related to HR repair. Another drug class identified was that of PI3K inhibitors targeting its oncogenic pathway. In line with these findings, across tumors a large fraction of patients show genomic profiles that could influence therapeutic choices besides current disease-specific indications. Recently, the MSK-IMPACT study[58] revealed that 34% of patients harbored at least one mutation classified as a predictive biomarker of response to an FDA-approved or investigational drug according to supportive clinical evidence. Though several issues should be considered such as intratumor heterogeneity as well as cancer specificity in drug response, our data indicate a space for drug repurposing strategies, whose benefit has been broadly reported across tumor types.[59,60] In TC, valuable data have been recently published on late-stage, heavily treated patients, for whom molecularly guided therapies led to a progression free-survival ratio of 1.4 with respect to non-molecularly guided treatment,[54] indicating potential advantages of targeted-based therapy. Notwithstanding that *ad hoc* prospective trials are needed to clearly assess the prognostic and predictive relevance of genetic biomarkers and validate the feasibility of molecularly informed therapies, the detection of extensive alterations in TC suggests that germline testing on a limited number of genes might be valuable to identify clinically relevant biomarkers, assessing, for instance, variants in a small panel of DDR genes such as core HR and MMR players *BRCA1*, *BRCA2*, *ATR*, *ATM*, *CHEK2*, *PALB2*, *MLH1*, *MSH2*, *MSH6*, and *PMS2*. In line with this idea, Möhrmann et al. recommend broad germline testing independent of clinical parameters for all patients with TC and high-grade thymomas based on their recent data.[54] Information on patients' family history was not available in this study, but our results highlight the importance of collecting data on family history to better evaluate cancer susceptibility in TC and the need for genetic counselling.

As for genomic imbalances, tumors displayed a highly heterogeneous degree of SCNAs, with few TCs characterized by the absence of SCNA and LOH events or by a high level of variation. The median SCNA burden of 0.23 was in line with the only report by Yang et al.,[46] whereas median LOH burden was firstly reported with a value of 0.11. In cancer, aneuploidy and high SCNA burden have been associated with aggressive tumor behavior and poor prognosis[61] in a variety of solid malignancies, including prostate, breast, endometrial, and colorectal cancers, likely mediated by an increase in tumor resistance to therapy and metastatic dissemination.[62] In this study, a high degree of SCNAs pointed to an association with advanced tumor stages. Although a negative trend of association with survival was detected for the number of SCNA and LOH events, further investigation is required to clearly assess their prognostic value as biomarkers. Tumors displayed recurrent SCNAs, of which amplifications at chromosomes 1q21.3, 7p14.1, and 7q34 and deletions at chromosomes 2q37.1, 9p21.3, and 14q11.2 were previously described in TCs and TETs, as well as broader losses at chromosomes 16q and 17q.[12,13] Notably, amplification at chromosomes 3q29, 8p21.2, and 14q11.2 was negatively associated with survival. Few data about genomic amplification at chromosome 3q29 are available in cancer. Genome-wide association studies linked this alteration with susceptibility to pancreatic, colorectal, gastric, and prostate cancers.[63,64] Somatic gains at 3q29 were described in B-cell lymphoma, lung cancer, and also in TC.[13,65,66] In this region, *MUC4* oncogene encodes a membrane-bound mucin, which has been involved in the regulation of processes related to metastasis and therapeutic resistance.[67] Although genomic imbalance at chromosome 8p21 was frequently observed as large-scale LOH events in tumors,[68] evidence of amplification at chromosome 8p in TC was also observed in other studies,[12,13] an aspect that might suggest tissue specificity in the selection of this aberration. Gains at chromosome 14q11.2 have rarely been reported in cancer. Sporadic gains have recently been described in oral squamous-cell carcinoma,[69] whereas amplification of regions 14q11.2-q21 and 14q11-q24 has been reported once in thyroid and breast cancer,[70,71] but no gene has been identified as a driver so far.

This work has several strengths. Firstly, the sequencing approach allowed the analysis of genomic alterations

besides the limited pool of cancer-related genes and the assessment of SCNAs. Secondly, the availability of matched-normal samples allowed the characterization of germline variants possibly related to tumor susceptibility. Thirdly, the presence of matched-normal tissues, along with unrelated control non-cancer samples, enhanced the confidence of genomic calls. This work also presents some limitations. Firstly, the small sample size of the study, though inherently related to tumor rarity, requires the validation of results with more precise estimation of effect sizes and control of *P* values in the assessment of the prognostic value of biomarkers. Secondly, due to the unavailability of matched-normal samples for a few tumors, rare germline variants might have been included in the somatic calling despite stringent filtering on allelic fraction. Lastly, the unavailability of multiple and longitudinal samples per patient impaired a mutational clonality analysis, which is essential to assess drivers of tumor evolution and evaluate the feasibility of biomarker-based therapeutic strategies.

FUNDING

This work was supported by the research organization Alleanza contro il Cancro [grant identifier RCR-2021-23671213—Cat.B-PRV-RETE ACC—WP 10 Genomica—Consolidamento di un programma multicentrico di genomica delle neoplasie rare tramite i progetti UNIC (Unmet Needs in Clinical Genomics); https://alleanzacontroilcancro.it/]. The publication fee for this work was covered by the Italian Ministry of Health's "Ricerca Corrente" funding to the IRCCS Humanitas Research Hospital.

DISCLOSURE

The authors have declared no conflicts of interest.

REFERENCES

1. de Jong WK, Blaauwgeers JLG, Schaapveld M, Timens W, Klinkenberg TJ, Groen HJM. Thymic epithelial tumours: a population-based study of the incidence, diagnostic procedures and therapy. *Eur J Cancer.* 2008;44(1):123-130.
2. Scorsetti M, Leo F, Trama A, et al. Thymoma and thymic carcinomas. *Crit Rev Oncol Hematol*. 2016;99:332-350.
3. Girard N, Ruffini E, Marx A, Faivre-Finn C, Peters S. Thymic epithelial tumours: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2015;26(suppl 5):v40-v55.
4. Zucali PA, de Vincenzo F, Perrino M, et al. Systemic treatments for thymic tumors: a narrative review. *Mediastinum*. 2021;5:24.
5. Sato J, Satouchi M, Itoh S, et al. Lenvatinib in patients with advanced or metastatic thymic carcinoma (REMORA): a multicentre, phase 2 trial. *Lancet Oncol*. 2020;21(6):843-850.
6. Thomas A, Rajan A, Berman A, et al. Sunitinib in patients with chemotherapy-refractory thymoma and thymic carcinoma: an open-label phase 2 trial. *Lancet Oncol*. 2015;16(2):177-186.
7. Zucali PA, De Pas T, Palmieri G, et al. Phase II study of everolimus in patients with thymoma and thymic carcinoma previously treated with cisplatin-based chemotherapy. *J Clinl Oncol*. 2018;36(4):342-349.
8. Giaccone G, Kim C, Thompson J, et al. Pembrolizumab in patients with thymic carcinoma: a single-arm, single-centre, phase 2 study. *Lancet Oncol*. 2018;19(3):347-355.
9. Pritchard CC, Mateo J, Walsh MF, et al. Inherited DNA-repair gene mutations in men with metastatic prostate cancer. *N Engl J Med*. 2016;375(5):443-453.
10. Robson M, Im SA, Senkus E, et al. Olaparib for metastatic breast cancer in patients with a germline BRCA mutation. *N Engl J Med*. 2017;377(6):523-533.
11. Wang Y, Thomas A, Lau C, et al. Mutations of epigenetic regulatory genes are common in thymic carcinomas. *Sci Rep*. 2014;4:7336.
12. Radovich M, Pickering CR, Felau I, et al. The integrated genomic landscape of thymic epithelial tumors. *Cancer Cell*. 2018;33(2):244-258.e10.
13. Petrini I, Meltzer PS, Zucali PA, et al. Copy number aberrations of BCL2 and CDKN2A/B identified by array-CGH in thymic epithelial tumors. *Cell Death Dis*. 2012;3(7):e351.
14. Repetto M, Conforti F, Pirola S, et al. Thymic carcinoma with Lynch syndrome or microsatellite instability, a rare entity responsive to immunotherapy. *Eur J Cancer.* 2021;153:162-167.
15. Katsuya Y, Fujita Y, Horinouchi H, Ohe Y, Watanabe S-I, Tsuta K. Immunohistochemical status of PD-L1 in thymoma and thymic carcinoma. *Lung Cancer.* 2015;88(2):154-159.
16. Marx A, Chan JKC, Chalabreysse L, et al. The 2021 WHO classification of tumors of the thymus and mediastinum: what is new in thymic epithelial, germ cell, and mesenchymal tumors? *J Thorac Oncol*. 2022;17(2):200-213.
17. Ruffini E, Fang W, Guerrera F, et al. The International Association for the Study of Lung Cancer Thymic Tumors Staging Project: the impact of the eighth edition of the Union for International Cancer Control and American Joint Committee on Cancer TNM Stage Classification of Thymic Tumors. *J Thorac Oncol*. 2020;15(3):436-447.
18. Mannucci PM, Merlini PA, Ardissino D, et al. No evidence of association between prothrombotic gene polymorphisms and the development of acute myocardial infarction at a young age. *Circulation*. 2003;107(8):1117-1122.
19. Do R, Stitziel NO, Won HH, et al. Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction. *Nature*. 2015;518(7537):102-106.
20. Ho DE, Imai K, King G, Stuart EA. MatchIt: nonparametric preprocessing for parametric causal inference. *J Stat Softw.* 2011;42(8):1-28.
21. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-1760.
22. McKenna A, Hanna M, Banks E, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297-1303.
23. Benjamin D, Sato T, Cibulskis K, Getz G, Stewart C, Lichtenstein L. Calling somatic SNVs and indels with Mutect2. *bioRxiv.* 2019. https://doi.org/10.1101/861054.
24. Kim S, Scheffler K, Halpern AL, et al. Strelka2: fast and accurate calling of germline and somatic variants. *Nat Methods*. 2018;15(8):591-594.
25. Lai Z, Markovets A, Ahdesmaki M, et al. VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. *Nucleic Acids Res*. 2016;44(11):e108.
26. Poplin R, Ruano-Rubio V, DePristo MA, et al. Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv.* 2017. https://doi.org/10.1101/201178.
27. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol*. 2016;17(1):122.
28. Ioannidis NM, Rothstein JH, Pejaver V, et al. REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. *Am J Hum Genet*. 2016;99(4):877-885.
29. Tamborero D, Rubio-Perez C, Deu-Pons J, et al. Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations. *Genome Med*. 2018;10(1):25.
30. Jia P, Yang X, Guo L, et al. MSIsensor-pro: fast, accurate, and matched-normal-sample-free detection of microsatellite instability. *Genomics Proteomics Bioinformatics*. 2020;18(1):65-71.
31. Abida W, Cheng ML, Armenia J, et al. Analysis of the prevalence of microsatellite instability in prostate cancer and response to immune checkpoint blockade. *JAMA Oncol*. 2019;5(4):471-478.

32. Riester M, Singh AP, Brannon AR, et al. PureCN: copy number calling and SNV classification using targeted short read sequencing. *Source Code Biol Med*. 2016;11(1):13.
33. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol*. 2011;12(4):R41.
34. Blokzijl F, Janssen R, van Boxtel R, Cuppen E. MutationalPatterns: comprehensive genome-wide analysis of mutational processes. *Genome Med*. 2018;10(1):33.
35. Bergstrom EN, Huang MN, Mahto U, et al. SigProfilerMatrixGenerator: a tool for visualizing and exploring patterns of small mutational events. *BMC Genomics*. 2019;20(1):685.
36. Sondka Z, Dhir NB, Carvalho-Silva D, et al. COSMIC: a curated database of somatic variants and clinical data for cancer. *Nucleic Acids Res*. 2024;52(D1):D1210-D1217.
37. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015;1(6):417-425.
38. Zhou Y, Zhou B, Pache L, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019;10(1):1523.
39. Wilcoxon F. Individual comparisons by ranking methods. *Biometrics Bull*. 1945;1(6):80-83.
40. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*. 1958;53(282):457-481.
41. Mantel N. Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemother Rep*. 1966;50(3):163-170.
42. Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. *Control Clin Trials*. 1996;17(4):343-346.
43. Cox DR. Regression models and life-tables. *J R Stat Soc Ser B (Methodol)*. 1972;34(2):187-202.
44. Landrum MJ, Lee JM, Benson M, et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res*. 2018;46(D1):D1062-D1067.
45. Girard N, Basse C, Schrock A, Ramkissoon S, Killian K, Ross JS. Comprehensive genomic profiling of 274 thymic epithelial tumors unveils oncogenic pathways and predictive biomarkers. *Oncologist*. 2022;27(11):919-929.
46. Yang J, Zhang B, Guan W, et al. Molecular genetic characteristics of thymic epithelial tumors with distinct histological subtypes. *Cancer Med*. 2023;12(9):10575-10586.
47. Kurokawa K, Shukuya T, Greenstein RA, et al. Genomic characterization of thymic epithelial tumors in a real-world dataset. *ESMO Open*. 2023;8(5):101627.
48. 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.
49. Choi J, Park KH, Kim YH, et al. Large-scale cancer genomic analysis reveals significant disparities between microsatellite instability and tumor mutational burden. *Cancer Epidemiol Biomarkers Prev*. 2024;33(5):712-720.
50. Vanderwalde A, Spetzler D, Xiao N, Gatalica Z, Marshall J. Microsatellite instability status determined by next-generation sequencing and compared with PD-L1 and tumor mutational burden in 11,348 patients. *Cancer Med*. 2018;7(3):746-756.
51. Wu HX, Wang ZX, Zhao Q, et al. Tumor mutational and indel burden: a systematic pan-cancer evaluation as prognostic biomarkers. *Ann Transl Med*. 2019;7(22):640.
52. Conforti F, Zucali PA, Pala L, et al. 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(10):1287-1296.
53. NCCN. *NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines) for Guideline Thymomas and Thymic Carcinomas version 1.2024. A National Comprehensive Cancer Network, Inc.*. All rights reserved Available at https://thymic.org/wp-content/uploads/2024/07/NCCN-Guidelines-Version-1.2024.pdf. Accessed March 1, 2026.
54. Möhrmann L, Rostock L, Werner M, et al. Genomic landscape and molecularly informed therapy in thymic carcinoma and other advanced thymic epithelial tumors. *Med*. 2025;6(6):100612.
55. Yap TA, Ashok A, Stoll J, et al. Prevalence of germline findings among tumors from cancer types lacking hereditary testing guidelines. *JAMA Netw Open*. 2022;5(5):e2213070.
56. Huang K-L, Mashl RJ, Wu Y, et al. Pathogenic germline variants in 10,389 adult cancers. *Cell*. 2018;173(2):355-370.e14.
57. Lincoln SE, Nussbaum RL, Kurian AW, et al. Yield and utility of germline testing following tumor sequencing in patients with cancer. *JAMA Netw Open*. 2020;3(10):e2019452.
58. Zehir A, Benayed R, Shah RH, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med*. 2017;23(6):703-713.
59. Sicklick JK, Kato S, Okamura R, et al. Molecular profiling of cancer patients enables personalized combination therapy: the I-PREDICT study. *Nat Med*. 2019;25(5):744-750.
60. Principe DR, Kamath SD, Munshi HG, Mohindra NA. Metastatic thymoma harboring a deleterious BRCA2 mutation derives durable clinical benefit from olaparib. *Oncologist*. 2020;25(4):301-305.
61. Hieronymus H, Murali R, Tin A, et al. Tumor copy number alteration burden is a pan-cancer prognostic factor associated with recurrence and death. *Elife*. 2018;7:e37294.
62. Tijhuis AE, Johnson SC, McClelland SE. The emerging links between chromosomal instability (CIN), metastasis, inflammation and tumour immunity. *Mol Cytogenet*. 2019;12(1):12-17.
63. Childs EJ, Mocci E, Campa D, et al. Common variation at 2p13.3, 3q29, 7p13 and 17q25.1 associated with susceptibility to pancreatic cancer. *Nat Genet*. 2015;47(8):911-916.
64. Samola Winnberg J, Vermani L, Liu W, et al. A genome-wide association study in Swedish colorectal cancer patients with gastric- and prostate cancer in relatives. *Hered Cancer Clin Pract*. 2024;22(1):25.
65. Kang JU, Koo SH, Kwon KC, Park JW, Kim JM. Identification of novel candidate target genes, including EPHB3, MASP1 and SST at 3q26.2-q29 in squamous cell carcinoma of the lung. *BMC Cancer*. 2009;9:237.
66. Grygalewicz B, Szafron LM, Szafron LA, et al. Cytogenomic and clinicopathologic comparison of MYC-positive and MYC-negative high-grade B-cell lymphoma with 11q aberration in the context of other aggressive lymphomas with MYC rearrangement. *Mod Pathol*. 2025;38(8):100774.
67. Dreyer CA, VanderVorst K, Free S, Rowson-Hodel A, Carraway KL. The role of membrane mucin MUC4 in breast cancer metastasis. *Endocr Relat Cancer*. 2022;29(1):R17-R32.
68. Eichenauer T, Bannenberg DC, Kluth M, et al. 8p deletions in renal cell carcinoma are associated with unfavorable tumor features and poor overall survival. *Urol Oncol*. 2020;38(2):43.e13-43.e20.
69. Chen TW, Lee CC, Liu H, et al. APOBEC3A is an oral cancer prognostic biomarker in Taiwanese carriers of an APOBEC deletion polymorphism. *Nat Commun*. 2017;8(1):465.
70. Zudaire I, Odero MD, Caballero C, et al. Genomic imbalances detected by comparative genomic hybridization are prognostic markers in invasive ductal breast carcinomas. *Histopathology*. 2002;40(6):547-555.
71. Chen X, Knauf JA, Gonsky R, et al. From amplification to gene in thyroid cancer: a high-resolution mapped bacterial-artificial-chromosome resource for cancer chromosome aberrations guides gene discovery after comparative genome hybridization. *Am J Hum Genet*. 1998;63(2):625-637.