

# Unraveling the Regulation of Cancer/Testis Antigens in Tumorigenesis Through an Analysis of Normal Germ Cell Development in Rodents



Haiqi Chen, Yu Jiang, Dolores D. Mruk, and C. Yan Cheng

## Introduction

During the past two decades, much effort has been made to enhance anti-tumor immunological responses in patients by tumor antigen-specific immunization, which targets cancer cells [1, 2]. Among the different types of tumor antigens, cancer/testis (CT) antigens are promising therapeutic targets for the development of cancer vaccines due to their unique characteristics. Under normal physiological conditions, CT antigens are predominantly expressed in germline cells, but also in spermatocytes, and even in elongating/elongated spermatids. Since the testis is an immune privileged organ and also an immune privileged site, indicating that foreign tissue grafts into the testis can survive longer, or indefinitely, without being rejected; or the testis also resists rejection when its tissue is grafted into a non-privileged site [3, 4]. CT antigens residing in germ cells do not provoke immune response, and thus avoiding their eradication by the immune system. Interestingly, CT antigens can be aberrantly expressed in multiple cancers, which is suggestive of an aberrant reactivation of the silenced gametogenic program in somatic cells during tumorigenesis [5, 6]. Thus, CT antigens hold great promise for highly specific and safe immunotherapy. The first CT antigen MAGE1 (melanoma-associated antigen 1) was

---

H. Chen · D. D. Mruk

The Mary M. Wohlford Laboratory for Male Contraceptive Research, Center for Biomedical Research, Population Council, New York, NY, USA

Y. Jiang

College of Pharmacy, University of Iowa, Iowa City, IA, USA

C. Y. Cheng (✉)

Sir Run Run Shaw Hospital (SRRSH), Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

e-mail: [yancheng01@aol.com](mailto:yancheng01@aol.com); [cyancheng@ntu.edu.cn](mailto:cyancheng@ntu.edu.cn)

discovered more than two decades ago through autologous typing [7]. Since then, more than 200 CT antigens have been identified [8]. Some of these CT antigens have been tested as tumor vaccines in clinical trials. Unfortunately, most clinical studies involving CT antigens have so far been unsuccessful [9–11], partly due to the fact that there is a lack of systematic understanding of the expression pattern of CT antigens across different cancer types, as well as the molecular mechanisms underlying their aberrant expression and oncogenic functions in cancer cells.

Previous studies have reported shared characteristics between germ cells and cancer cells (for a review, see [6]). For instance, CT antigens that are involved in germ cell implantation and migration during gametogenesis can regulate cancer cell invasion and metastasis [6, 12]. Therefore, it is likely that similar mechanisms underlie the regulation of CT antigens during gametogenesis and their ectopic expression during tumorigenesis. Thus, understanding the regulation of CT antigens during gametogenesis may provide valuable insights into their regulation in the context of cancers. Recently, one research group has carried out a comprehensive analysis using publicly available databases to systematically identify CT antigens, which has significantly expanded the repertoire of CT antigens [13]. Furthermore, the past few years have witnessed a significant surge in publicly available next generation sequencing datasets both from the fields of reproduction and cancer. In this review, we analyzed these datasets and focused our attention on the regulation of CT antigens in the contexts of gametogenesis and tumorigenesis. We found that many aberrantly expressed CT antigens in cancer cells associated with super-enhancers, and some of these associations were conserved in normal testes and cancers. We also reported that methyltransferase DNMT1 is a potential regulator of the expression of CT antigens, and identified CT antigen and transcription factor TAF7L as an important regulator of the expression of a subset of CT antigens. This information may be useful in the identification of new targets among CT antigens for the treatment of cancer. This review also provides a credible approach of using big datasets deposited in public domain for meaningful analysis by investigators in the field.

## Approaches Used for Analysis

1. **Data sources.** Raw datasets from published studies were retrieved from the Gene Expression Omnibus (GEO) (<https://www.ncbi.nlm.nih.gov/geo/>). A complete list of the GEO accession numbers used in this report is provided in Table 1. Gene expression and DNA methylation data (Illumina 450k platform) from the TCGA Database were retrieved via the UCSC Xena Browser (<http://xena.ucsc.edu/>). Cancer samples lacking expression of target CT antigens were removed from the analysis. Gene expression data from normal human tissues were obtained from the GTEx Database via the UCSC Genome Browser (<https://genome.ucsc.edu/>). Promoter information was obtained from the Eukaryotic Promoter Database (<http://epd.vital-it.ch/index.php>). CpG island information was acquired from the UCSC Genome Browser. The MiPanda Michigan Portal for the Analysis of NGS Data (<http://www.mipanda.org/>) and the Firehose

**Table 1** List of GEO accession numbers used in this report

Data name	GEO accession number	Data type
Expression profiling of mouse PGCs from DNMT1 KO model	GSE74938	Bulk RNA-Seq
Expression profiling of human PGCs	GSE63818	Single cell RNA-Seq
Chromatin IP against multiple factors and histone modifications in multiple human cancer cell lines	GSE36354	ChIP-Seq
Chromatin IP against H3K27ac in normal testis	GSE75734	ChIP-Seq
Chromatin IP against TAF7L and POLII in mouse testis; expression profiling of mouse testis from TAF7L KO model	GSE50807	ChIP-Seq; Bulk RNA-Seq
Expression profiling of mouse PGCs	GSE41908	Bulk RNA-Seq
Expression profiling of purified mouse spermatogonia, spermatocytes, spermatids and spermatozoa	GSE43717	Bulk RNA-Seq

Browser (<https://gdac.broadinstitute.org/>) were also used to search and validate the data during the preparation of this report.

2. **RNA-Seq analysis.** RNA-Seq analysis, bulk RNA-Seq or single cell (sc) RNA-Seq, was performed based on the ‘new tuxedo’ pipeline [14]. In brief, raw reads were aligned to the mouse reference genome mm10 or human genome hg38 using the HISAT2 Program (version 2.1.0) [15] with default parameters. BAM files were then sorted and converted into SAM files using SAM Tools (version 1.4) (<http://www.htslib.org/>). The expressed genes and transcripts were assembled and quantified using the StringTie Program (version 1.3.3) [16]. **Differential gene expression** analysis was performed using the R/Bioconductor package Ballgown (version 2.10.0) [17]. Genes with FPKM = 0 (no fragments per kilobase of transcript per million mapped reads) were removed from the analysis. The multiple testing errors were corrected using the false discovery rate (FDR). Genes with a  $\geq 2$ -fold difference in expression and a FDR  $< 5\%$  were considered differentially expressed. For studies in which processed data were available, the results were cross-referenced with those from in-house analysis, and the analyzed data from the original studies were only used if the results matched.
3. **ChIP-Seq analysis.** To build version mm10 of the mouse genome or version hg38 of the human genome with default parameters, ChIP-Seq datasets were aligned using the Bowtie 2 Program (version 2.3.2) [18] BAM files were sorted and converted into SAM files using SAM Tools. Peak calling was performed in MACS2 (version 2.1.1) [19] to identify regions of ChIP-Seq enrichment over the background. A  $p$ -value threshold of enrichment of 0.05 was implemented. ChIP peak annotation was performed using the R/Bioconductor package ChIPseeker (version 1.14.1) [20]. ChIP peak coverage was visualized within the Integrative Genomics Viewer Desktop Application (version 2.4) [21, 22]. For studies in which processed data were available, the results were cross-referenced with those from in-house analysis, and the analyzed data from the original studies were only used if the results matched. SEs (Super-Enhancers) were identified using the ROSE algorithm (version 0.1) with default parameters [23, 24]. The annotations of SE-associated genes based on the proximity to SEs in normal

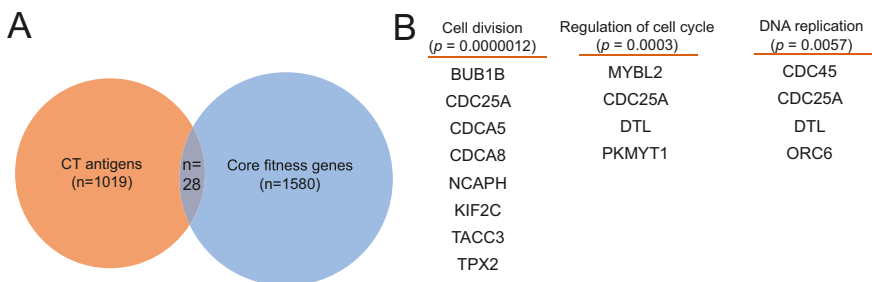
mouse testes were kindly provided by Dr. Alena Shkumatava from Institut Curie, PSL Research University.

4. **Functional annotations and statistical analysis.** Gene ontology analysis was performed using DAVID Online Tools (<http://david.abcc.ncifcrf.gov/>). The top-enriched clusters were identified by automated functional annotation clustering with default parameters.

Statistical analysis between two groups was performed using a non-parametric *t*-test. In all cases,  $p < 0.05$  was considered significant. Spearman's rank correlation coefficient ( $\rho$ ) was calculated as R (version 3.4.3) (<https://www.r-project.org/>).

## A Subset of CT Antigens Belongs to the Core Fitness Gene Family

We obtained the complete list of CT antigens systematically identified by a recent study in which 1019 testis-specific transcripts expressed in at least 1% of the samples from any of the 19 cancer types were examined [13]. We first asked if any of these CT antigens are human core fitness genes. The 1580 core fitness genes were defined as genes whose disruption decreased cell growth and proliferation in three or more of the five human cell lines (one normal immortalized cell line and four cancer cell lines) tested as revealed by CRISPR screens [25]. We compared the list of CT antigens with the list of human core fitness genes and identified 28 CT antigens belonging to the core fitness gene family (2.7% of the CT antigens and 1.8% of the core fitness genes) (Fig. 1a; Table 2). GO term enrichment analysis suggested that these 28 CT antigens were enriched in important cellular processes critical for cell survival such as cell division, cell cycle control and DNA replication (Fig. 1b). Therefore, these CT antigens may serve as potential targets for the treatment of cancer.



**Fig. 1** A subset of CT antigens belongs to the human core fitness gene family. (a) Venn diagram shows that 28 CT antigens belonged to the core fitness gene family whose perturbation may have decreased cell growth and proliferation. (b) These 28 CT antigens were enriched in cellular processes such as cell division, cell cycle control and DNA replication

**Table 2** CT antigens that belong to the family of core fitness genes

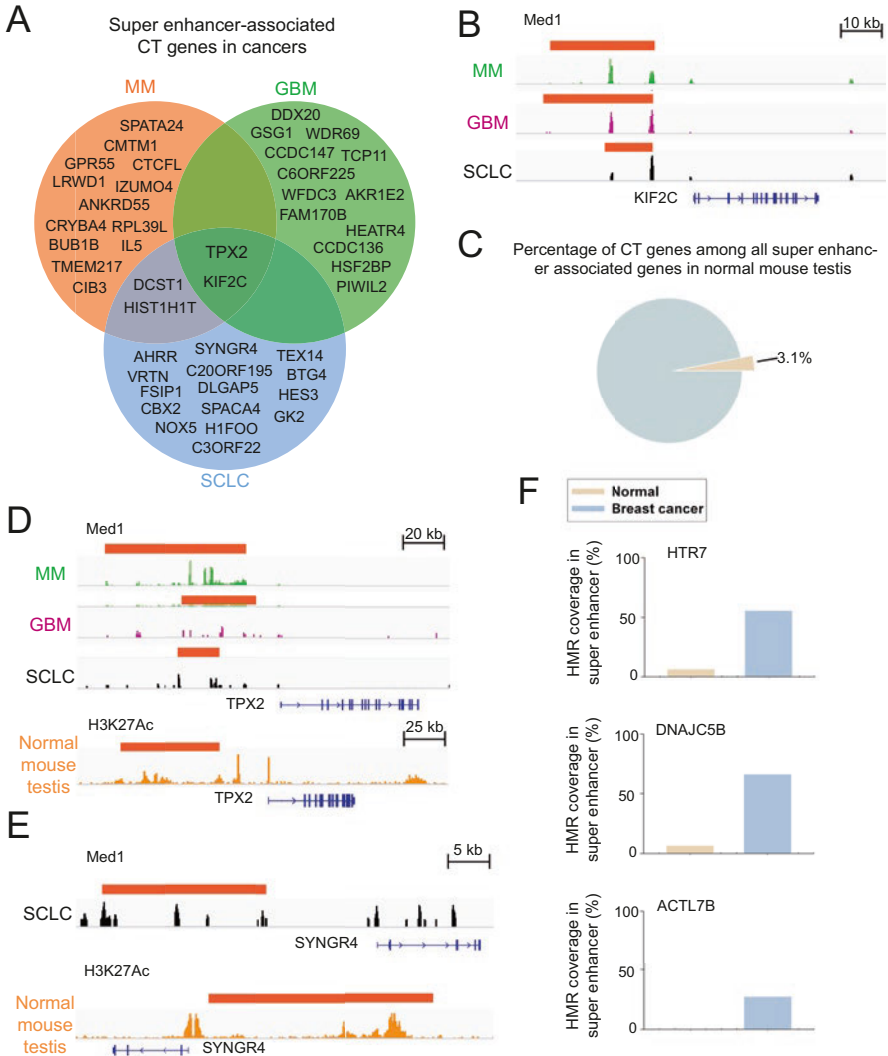
Gene names
PLK1
CT62
SGOL1
CDC25A
DDX20
BUB1B
RAD51
OIP5
MYBL2
CDC45
TPX2
ORC6
TTK
PKMYT1
TRIP13
HJURP
DMRTC2
CASC5
CDCA5
DTL
NUF2
KIF2C
CDCA8
KIF18A
NCAPH
TOP2A
PLK4
TACC3

## Multiple CT Antigens Are Under the Control of Super-Enhancers (SEs)

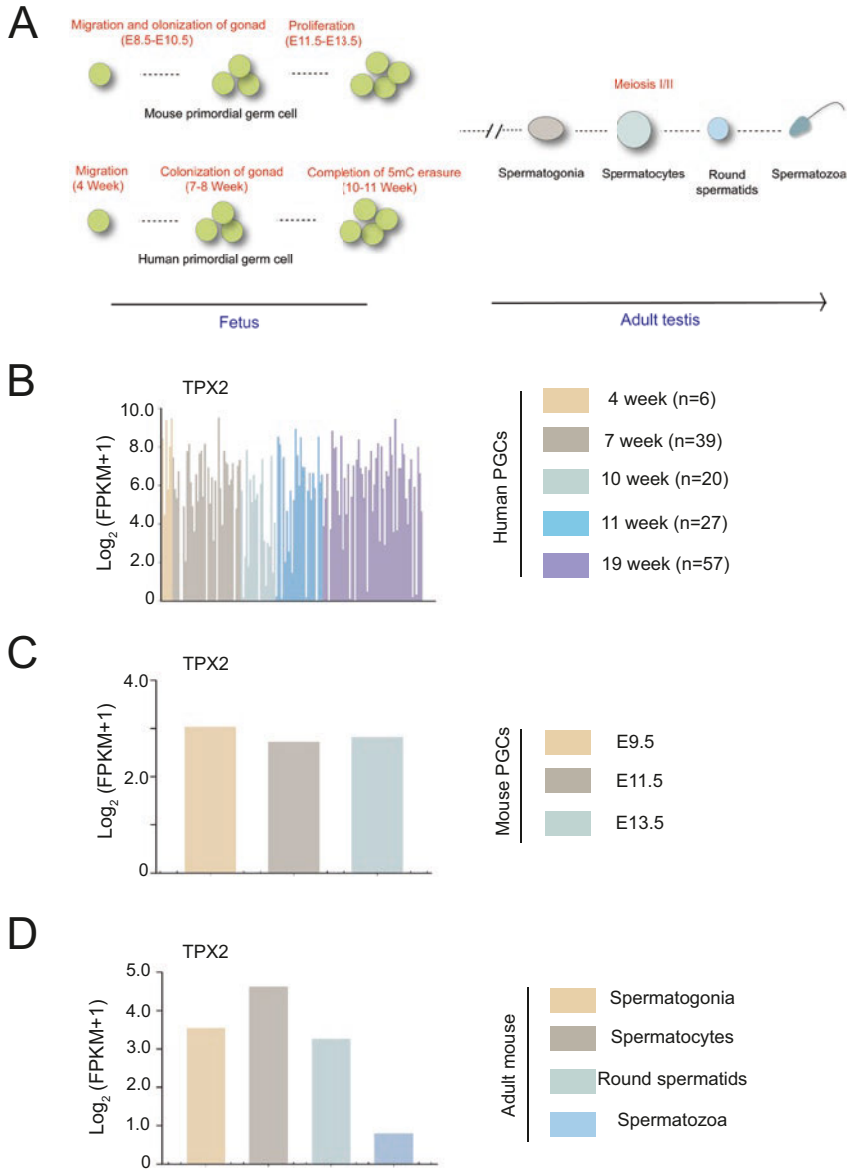
To gain insight into the regulation of CT antigens, we examined their association with SEs. SEs are large clusters of enhancers that are occupied by master transcription factors. They associate with key genes and are critical for cell identity [23]. Furthermore, SEs are hyperactive chromatin regions characterized by high levels of histone H3 lysine 27 acetylation (H3K27ac) that distinguish them from shorter, less active typical enhancer regions [23, 26]. SEs have been found at key oncogenic drivers in multiple cancer cells [24]. It has also been reported that cancer cells can acquire SEs at oncogenic drivers during the process of tumorigenesis as revealed by a comparison of SEs in cancer cells with corresponding normal cells [27]. Thus, SEs may be helpful for identifying key oncogenes in specific cancers. We accessed

the datasets for the lists of genes that associated with SEs in a multiple myeloma (MM) cell line, a glioblastoma multiforme (GBM) cell line and a small cell lung cancer (SCLC) cell line [24, 27] and looked for CT antigens in each gene list (Fig. 2). We identified 17, 15 and 19 CT antigens in MM, GBM and SCLC, respectively (Fig. 2a). Each cancer cell type contained a set of CT antigens that was largely different from the other types with only a few overlaps (four between MM and SCLC, two between MM and GBM, two between GBM and SCLC) (Fig. 2a). These results were consistent with earlier observations that SEs are usually cell-type-specific and play key roles in the control of cell identity [23]. Nonetheless, two SE-associated CT antigens, namely, TPX2 and KIF2C, were aberrantly expressed in the three cancer cell lines. The SEs associating with these two CT antigens were found at similar chromatin positions across the three cancer cell types, although the lengths of the SEs varied (Fig. 2b, d). TPX2 is a spindle-associated microtubule (MT)-binding protein [28]. It is required for MT assembly to facilitate the formation of the spindle during mitosis in HeLa cells [29] and meiosis in *Xenopus* eggs [30, 31]. To understand the role of TPX2 during gametogenesis, we accessed and analyzed RNA sequencing (RNA-Seq) datasets from human primordial germ cells (PGCs) [32], mouse PGCs [33, 34] and adult mouse germ cells at different developmental stages [35]. High expression of TPX2 was found throughout PGC development in both the human and the mouse, which is characterized by active cell division (Fig. 3a–c). High expression of TPX2 was also observed at the transition of adult mouse spermatocytes to round spermatids during which meiosis I/II takes place (Fig. 3a, d). This expression pattern of TPX2 during gametogenesis coincides with its reported functions. Therefore, it is plausible that the acquisition of SEs drives the aberrant expression of TPX2 in cancer cells, which facilitates cell division via mitosis. KIF2C belongs to the kinesin-13 family, which induces MT depolymerization to regulate chromosome movement and segregation during mitosis [36, 37]. It has been reported that KIF2C is overexpressed in colorectal carcinoma as well as several types of pancreatic, gastric, breast and head and neck cancers [38]. KIF2C is also overexpressed in gliomas, and it has been proposed to be a potential prognostic marker for glioma [39]. Thus, the acquisition of SEs may underlie the ectopic expression of KIF2C in different cancer types. Of interest, both TPX2 and KIF2C belong to the human core fitness gene family (Table 2). Taken collectively, our observations suggest that both TPX2 and KIF2C are potential targets for the treatment of cancer.

To further understand the association between SEs and CT antigens, we examined SE-associated CT antigens in normal testes. Due to a lack of ChIP-Seq (chromatin immunoprecipitation followed by high-throughput sequencing) data on H3K27ac in human testes, we obtained and analyzed the H3K27ac ChIP-Seq dataset from mouse adult testes [26]. Among all SE-associated genes, 164 CT antigens were identified, which accounted for 16.1% of the total number of CT antigens and 3.1% of all SE-associated genes in adult mouse testes (Fig. 2c). We then compared the list of SE-associated CT antigens in normal mouse testes with the list of SE-associated CT antigens identified in MM, GBM and SCLC cell lines (Fig. 2a). Ten SE-associated CT antigens were found in both lists. For instance, SYNGR4 belongs



**Fig. 2** A subset of CT antigens is under the control of super-enhancers (SEs). (a) Venn diagram shows SE-associated CT antigens in multiple myeloma (MM), glioblastoma (GBM) and small cell lung cancer (SCLC) cell lines, respectively. TPX2 and KIF2C are the two common CT antigens found in the three cancers. (b) CT antigen KIF2C associates with a SE (red rectangle) in MM, GBM and SCLC, respectively, as indicated by the chromatin immunoprecipitation (ChIP) peak coverage of Med1. (c) Pie chart shows that SE-associated CT antigens account for 3.1% of all SE-associated genes in normal mouse testes. (d) CT antigen TPX2 associates with a SE (red rectangle) in MM, GBM, SCLC and normal mouse testes, respectively, as indicated by ChIP peak coverage of Med1 and H3K27Ac. (e) CT antigen SYNGR4 associates with a SE (red rectangle) in both SCLC and normal mouse testes as indicated by ChIP peak coverage of Med1 and H3K27Ac. (f) Bar graphs show that the hypomethylated region (HMR) coverage of SE-associated CT antigens HTR7, DNAJC5B and ACTL7B, respectively, increases in breast cancer compared with normal tissues



**Fig. 3** CT antigen TPX2 is expressed throughout male germ cell development. **(a)** Schematic drawing shows the entire process of male germ cell development from early primordial germ cells (PGCs) to mature spermatozoa both in the mouse and human. Key biological events are highlighted in red. **(b)** Bar graph shows the expression of CT antigen TPX2 in human PGCs. FPKM, fragments per kilobase of transcript per million mapped reads. **(c)** Bar graph shows the expression of CT antigen TPX2 in mouse PGCs. **(d)** Bar graph shows the expression of CT antigen TPX2 in mouse developing male germ cells in testes



to the synaptogyrin family and functions in pulmonary carcinogenesis and tumor progression [40, 41]. SYNGR4 associated with a SE in both normal mouse testes and a SCLC cell line at similar locations (Fig. 2e). We also found that TPX2 (but not KIF2C) associated with a SE in normal mouse testes, and this TPX2-associated SE shared a similar chromatin location with that in MM, GBM and SCLC cell lines (Fig. 2d). Taken collectively, our findings show that the associations with SEs in a certain subset of CT antigens are conserved from normal testes to various cancers, suggesting that a conserved mechanism underlies the regulation of these CT antigens between gametogenesis and tumorigenesis.

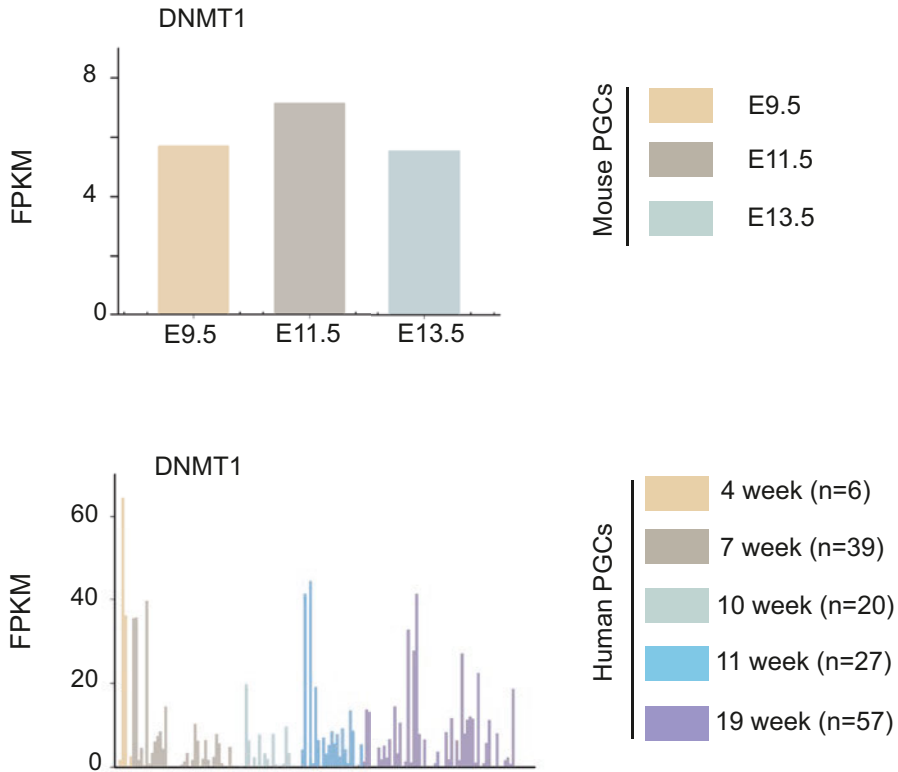
It was noted that only a small fraction of the associations (22.2% in our case) between the CT antigens and SEs were conserved from normal testes to cancers. This may have been due to the fact that mouse testes instead of human testes were used for the analysis. However, it may also have been due to the fact that there are other mechanisms responsible for the acquisition of SEs for CT antigens during tumorigenesis. The association of DNA methylation with enhancer elements has been proposed to influence enhancer activity [42, 43]. Thus, we speculated that an alteration in DNA methylation in SEs might contribute to the ectopic expression of certain CT antigens during tumorigenesis. To test this hypothesis, we obtained data from a recent study in which genome-scale DNA methylation analysis of normal tissues and multiple tumors were performed and the DNA methylation landscape across SEs among different normal tissue-matched tumors was reported [44]. We focused on breast cancers and found that three CT antigens (HTR7, DNAJC5B, ACTL7B) associating with SEs exhibited marked hypomethylation compared with normal tissues in which the occupancy of the hypomethylated regions (HMRs) increased dramatically in SE regions in the breast cancer samples (Fig. 2f). These results suggest that aberrant DNA methylation in SEs may be another contributing factor to the ectopic expression of CT antigens in cancers.

## **DNA Methyltransferase DNMT1 Is Involved in the Regulation of CT Antigen Expression**

The mission to identify the link between DNA methylation in SEs and CT antigen expression in different cancer types prompted us to further investigate the relationship between DNA methylation and CT antigen expression. The status of DNA methylation in selective genomic regions can be instructive for gene expression [43, 45]. In normal testes, aging-associated alterations in DNA methylation during spermatogenesis have been reported to have broad effects on gene expression and cell signaling in offspring [46]. During embryonic development, early PGCs undergo epigenetic reprogramming where their genomes display very low levels of global DNA methylation [47]. However, certain genomic regions, including CpG islands (regions where CpG is present at significantly higher levels than is typical for the genome as a whole) on the X chromosome and germline-specific genes, are

protected from the global loss of methylation, and they can only become demethylated upon gonad colonization of PGCs [48]. This protective effect has been shown to be mediated by methyltransferase DNMT1, because DNMT1 is the only DNA methyltransferase consistently expressed throughout the development of PGCs [49] (Fig. 4) and a DNMT1 knockout (KO) resulted in the up-regulation of post-migratory germ cell-specific genes [50]. Furthermore, a conditional loss of DNMT1 in PGCs leads to precocious differentiation in male germ cells by invoking cell lineage-specific gene expression [51]. Thus, we speculated that DNMT1 may also regulate the expression of specific CT antigens during PGC development. To test this hypothesis, we analyzed RNA-Seq data obtained from both wild type and DNMT1 KO male PGCs at E13.5 [51]. We found 141 differentially expressed genes with a fold change  $\geq 2$ . Among them, we identified 12 CT antigens that were differentially expressed in PGCs from the DNMT1 KO (Fig. 5a, b). Interestingly, all 12 differentially expressed CT antigens were up-regulated (Fig. 5a), which accounted for 11% of the total number of up-regulated genes (Fig. 5c). These results indicated that DNMT1 can suppress the expression of these CT antigens by maintaining DNA methylation during early PGC development. To confirm this, we looked at the methylation coverage of the promoter regions of these CT antigens in PGCs at E9.5 [48, 51]. Eight out of 12 CT antigens were found to have high methylation occupancy at their promoter regions ( $>20\%$ ) in PGCs at E9.5 (Fig. 5d), suggesting that the expression of these CT antigens may be suppressed by DNMT1 at early stages of PGC development.

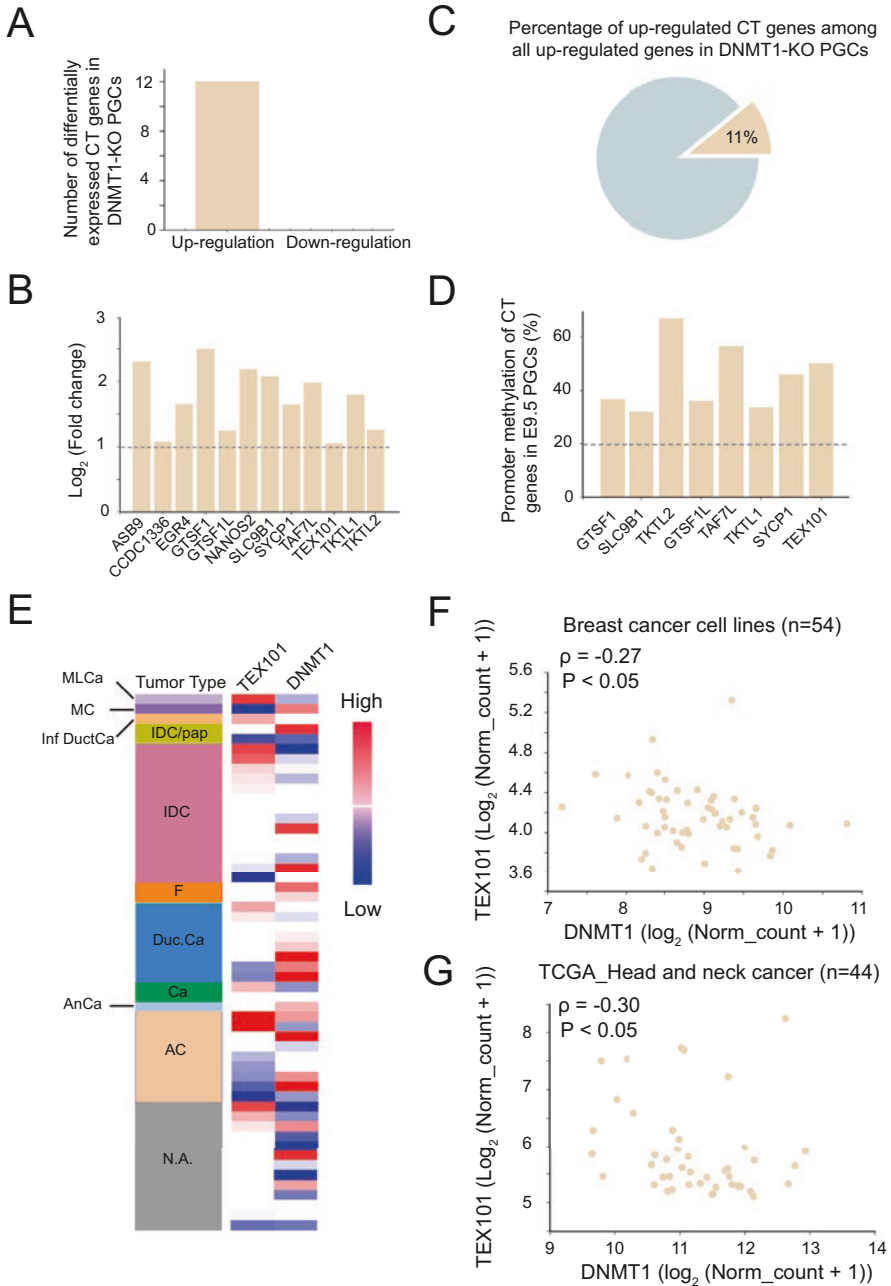
We then asked if a similar regulatory mechanism exists between DNMT1 and CT antigens in cancers. To answer this question, we focused on one of the eight CT antigens shown in Fig. 5d, TEX101, because it has been shown to play crucial roles in both spermatogenesis and tumorigenesis. TEX101 KO male mice were infertile due to failed uterotubal junction migration of their sperm [52, 53]. Its aberrant elevated expression has been reported in basal cell carcinoma [54], breast cancer [55, 56] and head and neck squamous cell carcinoma [57]. We first probed the expression levels of both TEX101 and DNMT1 across 54 breast cancer cell lines [58]. Indeed, TEX101 was overexpressed in specific breast cancer cell lines (Fig. 5e). An inverse correlation (Spearman's correlation, which applied to all correlation analyses in this report) was found between the expression levels of TEX101 and DNMT1 in these breast cancer cell lines (Fig. 5f), consistent with our observations from the PGC analysis. We also acquired gene expression data on TEX101 and DNMT1 from The Cancer Genome Atlas (TCGA) database. Interestingly, an inverse correlation between the expression of TEX101 and DNMT1 was also observed in head and neck cancer samples (Fig. 5g). Taken collectively, this information suggests that DNMT1 may negatively regulate the expression of these CT antigens in different cancer types.



**Fig. 4** DNA methyltransferase DNMT1 is expressed throughout primordial germ cell (PGC) development. Bar graphs show the expression of DNMT1 in both mouse and human PGCs

### The Expression of CT Antigen TAF7L Is Regulated by DNA Methylation in Cancers

We earlier showed that DNA methylation maintained by DNMT1 may control the expression of specific CT antigens during PGC development (Fig. 5b, d), and this may also hold true in cancer settings (Fig. 5e–g). To further test this idea in cancers, we focused on TAF7L, another CT antigen shown in Fig. 5d. TAF7L is a germ cell-specific transcription factor [59, 60]. Although TAF7L KO mice are fertile, their sperm exhibit abnormal morphology and significantly reduced motility [61]. Little is known about the role of TAF7L in cancers, although one study has reported a down-regulation of TAF7L expression in acute myeloid leukemia [62]. Two alternative promoter regions have been identified for TAF7L. The first promoter (promoter 1) was located at hg19 chrX: 100,548,043-100,548,102, whereas the second one (promoter 2) was located at hg19 chrX: 100,546,312-100,546,371 (Fig. 6a). A 5' CpG island was present at hg19 chrX: 100,546,064-100,546,550, associating



**Fig. 5** The expression of a subset of CT antigens is under the control of DNA methyltransferase DNMT1. (a) Bar graph shows that 12 CT antigens were found to be significantly up-regulated upon conditional deletion of DNMT1 in mouse primordial germ cells (PGCs) at E13.5. No CT antigens were significantly down-regulated. (b) Bar graph shows that the 12 up-regulated CT

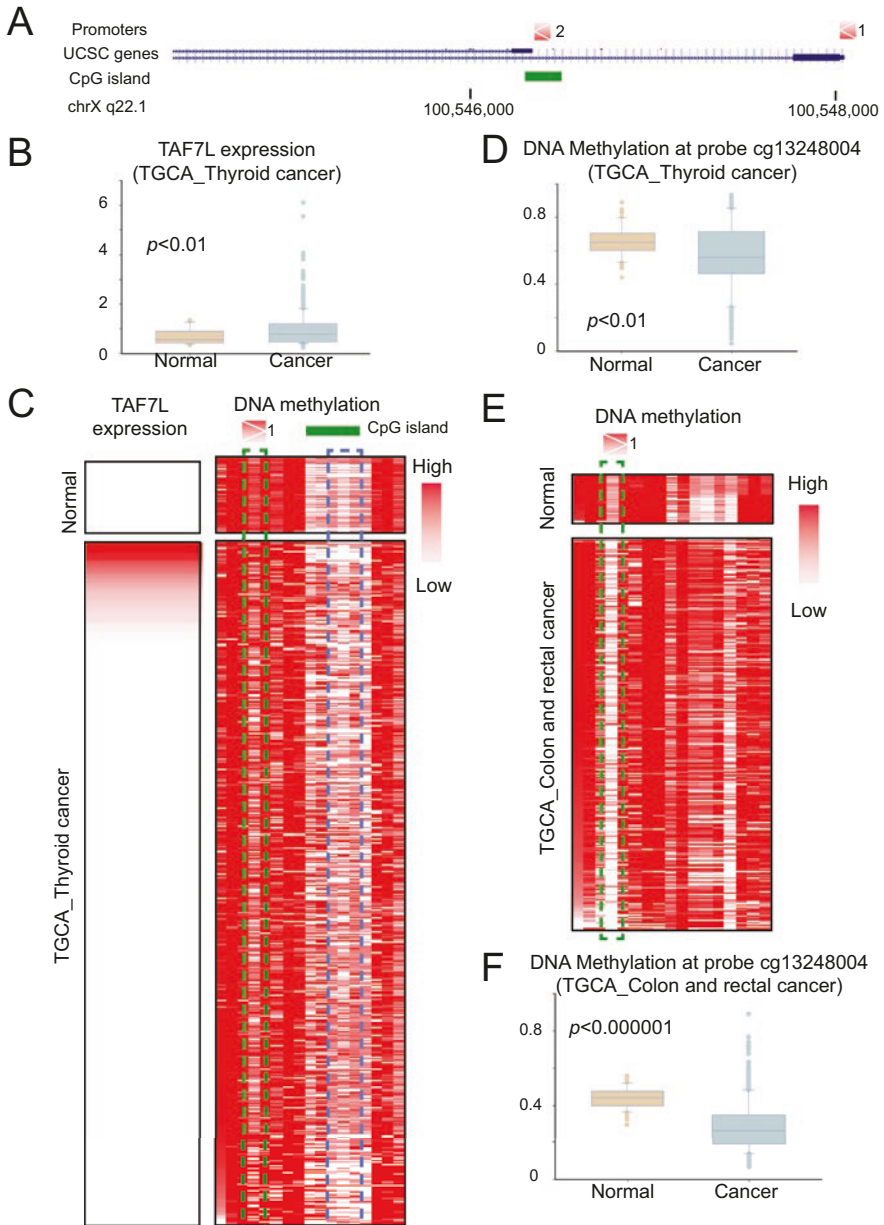
with promoter 2 of TAF7L (Fig. 6a). Furthermore, elevated expression of TAF7L was observed in thyroid cancers compared with normal tissues by analyzing TCGA datasets (Fig. 6b). We then set out to investigate if DNA methylation in the promoter regions of TAF7L or its associated CpG island affects its expression in thyroid cancer. We obtained DNA methylation profiles (Illumina 450k) of TAF7L in both normal and thyroid cancer tissues and coupled them with the TAF7L gene expression data (Fig. 6c). Interestingly, at probe cg13248004 (hg19 chrX: 100,548,042-100,548,044) near promoter 1 of TAF7L (Fig. 6c, **green box**), DNA methylation significantly decreased in cancer tissues compared with normal tissues (Fig. 6d). DNA demethylation at the same region was also observed in colon and rectal cancers (Fig. 6e, **green box** and Fig. 6f), suggesting that promoter 1 of TAF7L may undergo DNA demethylation during tumorigenesis. On the other hand, we observed an inverse correlation between the expression level of TAF7L and DNA methylation at probes cg08704539 (hg19 chrX: 100,546,355-100,546,357), cg01538344 (hg19 chrX: 100,546,290-100,546,292) and cg19386336 (hg19 chrX: 100,546,063-100,546,065) within the 5' CpG island in thyroid cancer tissues, respectively (Fig. 6c, **blue box**; Fig. 7). These findings indicate that DNA methylation at the promoter or CpG island plays a significant role in the ectopic expression of TAF7L in cancer.

## TAF7L Regulates the Expression of Other CT Antigens

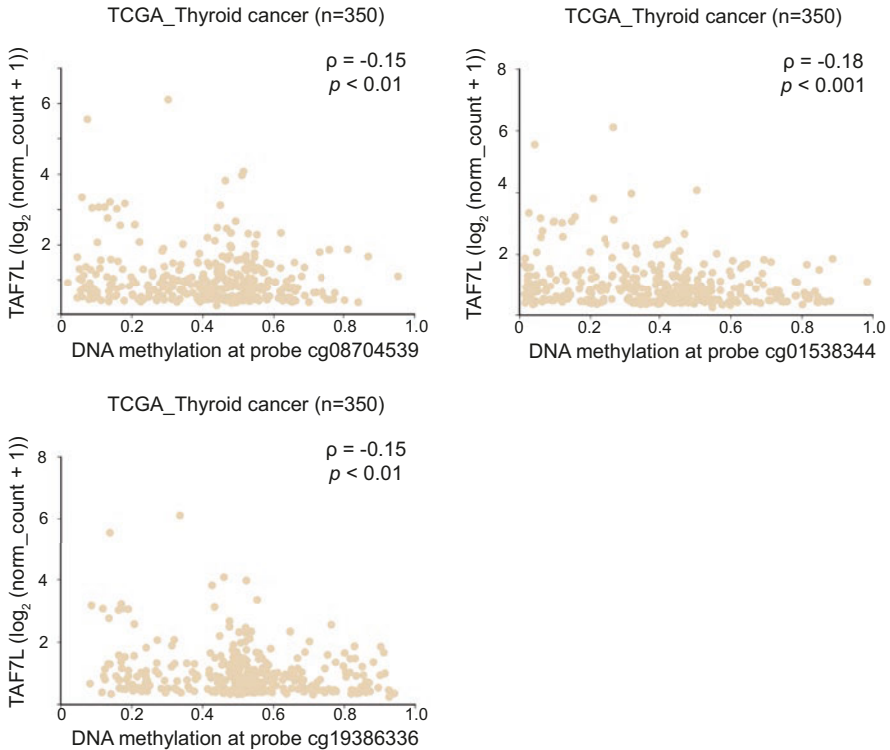
The fact that TAF7L is a transcription factor indicates that it can potentially regulate the expression of many other genes upon its activation during both gametogenesis and tumorigenesis. A recent study has shown that TAF7L targets promoters of many spermatogenesis-specific genes in testes [63]. We analyzed RNA-Seq data obtained from normal and TAF7L KO testes in this previous study and identified 267 differentially expressed CT antigens (fold change  $\geq 2$ ) upon TAF7L deletion (Fig. 8a-g). Among these differentially expressed CT antigens, 84 were up-regulated and 183 were down-regulated (Fig. 8a). We subsequently focused on the CT antigens that were down-regulated after TAF7L KO, because they accounted for the majority of the differentially expressed CT antigens. We filtered the 183 down-regulated CT



**Fig. 5** (continued) antigens showed  $>2$ -fold change in their expression levels. (c) Pie chart shows that the 12 up-regulated CT antigens account for 11% of all up-regulated genes in PGCs at E13.5. (d) Bar graph shows that eight of the 12 up-regulated CT antigen contained promoter regions with high DNA methylation (more than 20%) in PGCs at E9.5. (e) Heat map shows the expression of CT antigen TEX101, together with that of DNMT1, in different breast cancer cell lines. MLCa, metastatic lobular carcinoma; MC, metaplastic carcinoma; Inf DuctCa, infiltrating ductal carcinoma; IDC/pap, ductal breast cancer/papillary; IDC, ductal breast cancer; F, fibrocystic disease; Duc.Ca, ductal carcinoma; Ca, carcinoma; AnCa, anaplastic carcinoma; AC, adenocarcinoma; N.A., subtype identity not available. (f) Dot plot shows that the expression of CT antigen TEX101 inversely correlates with DNMT1 in breast cancer. (g) Dot plot shows that the expression of CT antigen TEX101 inversely correlates with DNMT1 in head and neck cancer



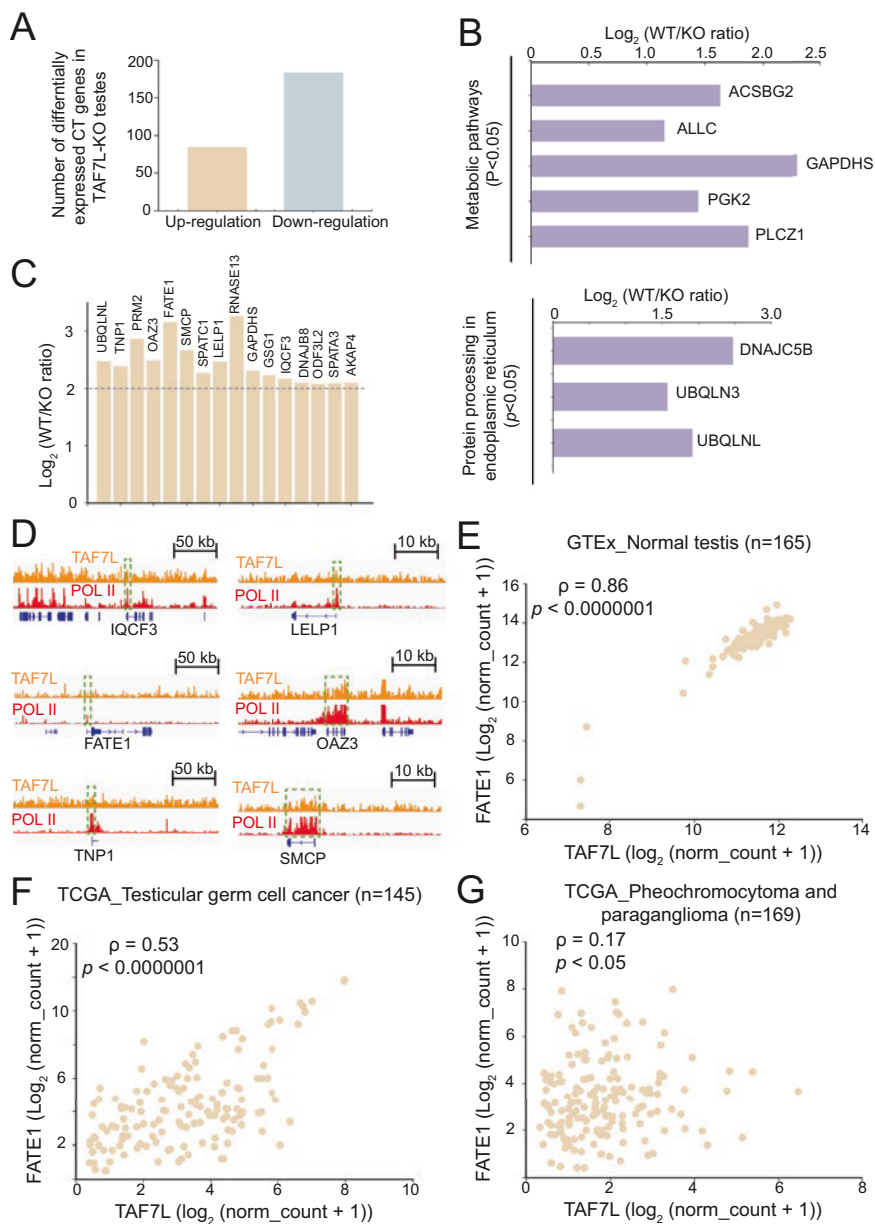
**Fig. 6** The expression of CT antigen TAF7L is under the control of DNA methylation in its promoter and 5' CpG island. (a) View of CT antigen TAF7L in UCSC genome browser. TAF7L has two alternative promoters. The 5' CpG island associated with promoter 2. (b) Box plot shows that the expression of TAF7L is significantly elevated in thyroid cancer compared with normal tissues. (c) Heat map shows the expression of TAF7L in both normal and thyroid cancer together with the DNA methylation status of the 5' region of TAF7L. (d) Box plot shows that promoter 1 (green box in c) of TAF7L is significantly hypomethylated in thyroid cancer compared with normal tissue. (e) Heat map shows the DNA methylation status of the 5' region of TAF7L. (f) Box plot shows that promoter 1 (green box in e) of TAF7L is significantly hypomethylated in colon and rectal cancer compared with normal tissues



**Fig. 7** The expression of CT antigen TAF7L is inversely correlated with DNA methylation of the 5' CpG island in thyroid cancer. Dot plots show that the expression of TAF7L negatively correlates with the DNA methylation status of its associated 5' CpG island in thyroid cancer. The exact location of each probe is described in the results session

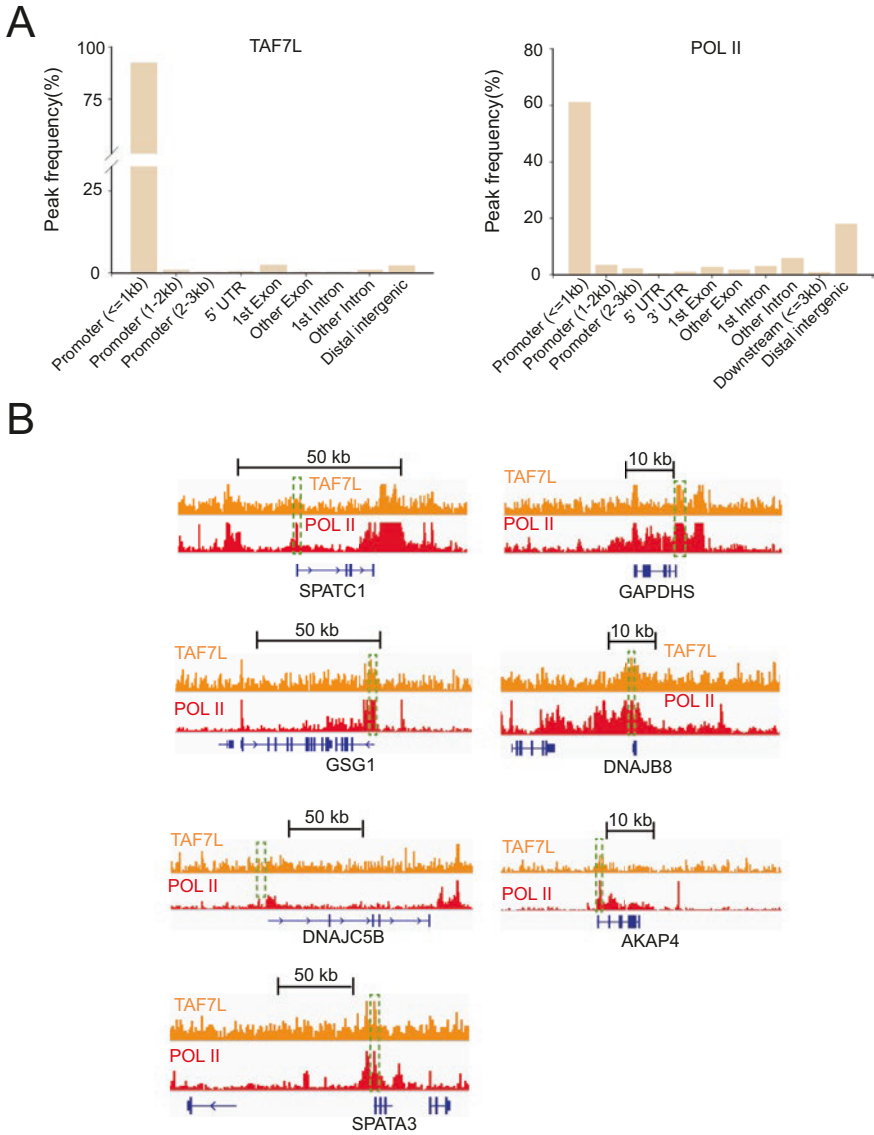
antigens by setting the threshold of their expression in normal testes as RPKM (reads per kilobase of transcript per million mapped reads)  $\geq 20$  in the GTEx (The Genotype-Tissue Expression) Database to rule out CT antigens whose expression levels were already low in normal testes. This resulted in a list of 103 CT antigens. Gene ontology analysis suggested that several CT antigens were enriched in protein processing and metabolic pathways (Fig. 8b), both of which are important in tumorigenesis. We then further filtered the 103 CT antigens by setting the threshold of fold change in their expression at 4 with the idea that the remaining CT antigens should be mostly susceptible to regulation by TAF7L. Sixteen CT antigens passed the threshold of fold change (Fig. 8c). We then confirmed if they were under direct control of TAF7L by analyzing the ChIP-Seq data on TAF7L and RNA polymerase II (POL II) [63]. Peak annotations showed that approximately 94% of TAF7L peaks were within 1 kb of the promoters (Fig. 9a), consistent with a previous analysis [63]. Moreover, the 16 filtered CT antigens were occupied by TAF7L peaks, and these CT antigens were all actively transcribed as indicated by the co-localized POL II peaks





**Fig. 8** CT antigen TAF7L controls the expression of a subset of CT antigens. **(a)** Bar graph shows that a total of 84 up-regulated and 183 down-regulated CT antigens are among the genes that were differentially expressed in TAF7L-deficient testes. **(b)** Down-regulated CT antigens identified in the TAF7L-deficient model were subject to functional annotations. Protein processing in the endoplasmic reticulum and metabolic pathways were enriched in the analysis. The fold change of each CT antigen is indicated. **(c)** Bar graph shows CT antigens with  $>4$ -fold change in expression identified in the TAF7L KO model. **(d)** CHIP peak coverage of TAF7L and POL II suggests that TAF7L directly binds to CT antigens, which are actively transcribed as indicated by colocalized peaks of Fig. 8 (continued) TAF7L and POL II (green box). **(e)** Dot plot shows that the expression of CT antigen FATE1 significantly correlates with the expression of TAF7L in normal human testes. **(f)** Dot plot shows that the expression of CT antigen FATE1 significantly correlates with the expression of TAF7L in testicular germ cell cancer. **(g)** Dot plot shows that the expression of CT antigen FATE1 significantly correlates with the expression of TAF7L in pheochromocytoma and paraganglioma





**Fig. 9** CT antigen TAF7L regulates the expression of a subset of CT antigens. **(a)** Bar graphs summarize peak annotations of the ChIP-Seq data for both TAF7L and POL II. **(b)** ChIP peak coverage of TAF7L and POL II indicates that TAF7L directly binds to CT antigens, which are actively transcribed as indicated by the colocalized peaks of TAF7L and POL II (green box)

(Figs. 8d and 9b, **green box**). These data suggest that TAF7L regulates the expression of these CT antigens by directly binding to their promoters in testes.

We then set out to test if a similar regulatory mechanism between TAF7L and CT antigens is found in cancers. We focused on one of the CT antigens FATE1 (Fig. 8c). FATE1 has been reported to be overexpressed in hepatocellular carcinoma, as well as gastric and colon cancers [64]. Furthermore, FATE1 decreases the sensitivity of adrenocortical carcinoma cells to the chemotherapeutic drug mitotane, and its expression in adrenocortical carcinoma negatively correlates with patient survival [65]. We observed that FATE1 expression was highly correlated with TAF7L expression in normal human testes (Fig. 8e), consistent with our earlier findings (Fig. 8c). Interestingly, this positive correlation persisted in testicular germ cell cancer (Fig. 8f), as well as pheochromocytoma and paraganglioma (Fig. 8g), indicating that FATE1 expression is also under the influence of TAF7L in cancer.

## Discussion and Concluding Remarks

The past decade has seen a tremendous surge in the number of high resolution next generation sequencing datasets generated from the fields of reproduction and cancer. New bioinformatics tools have also helped to systematically identify CT antigens, which have greatly expanded the pool of CT antigens from more than 200 to approximately 1000 [13]. At the time of this writing, a new study has reported yet another genome-wide identification of CT antigens, adding an additional 201 CT antigens to an already exhaustive pool of CT antigens [66]. With such an explosion of publicly available datasets, one may be surprised by how little we know about the molecular mechanisms underlying the function and regulation of these CT antigens throughout gametogenesis and tumorigenesis. The problem may no longer be a lack of available data to work on, but a lack of direction on where to look and how to analyze the data. The shared characteristics between gametogenesis and tumorigenesis have attracted considerable interest. It has been proposed that the expression of CT antigens during tumorigenesis is due to the reactivation of a gametogenic program in non-germline somatic cells [6, 12, 67]. If this is true, understanding the regulation of CT antigens during gametogenesis is likely to provide new insights on the regulation of these CT antigens during tumorigenesis.

By re-analyzing published datasets and data from multiple databases, we have re-examined the likely roles of CT antigens in cancers by providing the following information. First, a subset of CT antigens associated with SEs in three different cancer types. Of interest, ten of these CT antigens also associated with SEs in normal mouse testes. By analyzing ChIP-Seq data from human testes and several types of cancer, we expected to find additional conserved SE-associated CT antigens. Thus, it is tempting to speculate that these CT antigens acquired SEs via a conserved mechanism during tumorigenesis, which in turn drives their aberrant

expression in different cancer types. SEs have been reported to associate with key oncogenes [27, 68, 69]. We found that the CT antigen TPX2 associated with a SE in MM, GBM, and SCLC, as well as in normal testes. Moreover, TPX2 per se belongs to the human core fitness gene family. Previous studies have shown that certain inhibitors, such as the BET-bromodomain inhibitor JQ1 and CDK7 inhibitor, preferentially target SE-associated oncogenes in cancers [24, 70]. Thus, targeting TPX2 or its associated SE in cancer cells may prove to be an effective therapeutic approach.

Second, we showed that hypomethylation in SE regions may affect the ectopic expression of certain CT antigens in breast cancer. Hypomethylation/demethylation is known to occur throughout gametogenesis and tumorigenesis [47, 71]. Moreover, DNA hypomethylation participates in the activation of a subset of CT antigens in different cancer types [72, 73]. By extracting information from a study in which DNA methyltransferase DNMT1 was conditionally knocked out in PGCs, we found that DNMT1 can potentially suppress the expression of a new subset of CT antigens in cancers, such as TEX101, whose expression strongly correlates with cancer prognosis [66]. The demethylating agent 5-aza-2'-deoxycytidine has been shown to induce the expression of the CT antigen MAGE in both normal and malignant lymphoid cells [74]. In fact, 5-aza-2'-deoxycytidine inhibits DNA methylation by prohibiting DNMT1 access to the progressing replication fork during DNA replication [75]. The ablation of DNMT1 alone, but not DNMT3B alone, in HCT116 cells (a colorectal cancer cell line) activated CT antigens MAGE-A1, NY-ESO-1 and XAGE-1 [76]. Taken collectively, this information confirms our observations on the important role of DNMT1 in regulating CT antigen expression in cancer. Additionally, TAF7L is another CT antigen that was identified in the DNMT1 KO model, wherein the expression of TAF7L regulated by the status of DNA methylation in its promoter and the promoter-associated CpG island in cancers. This finding is not surprising given that demethylation of the CpG sites within the 5' regions of the CT antigen MAGE-A1 also correlates with its ectopic expression in cancer cells [77]. Thus, DNA hypomethylation may represent the primary mechanism of activation for a growing subset of CT antigens in cancers.

Finally, we examined a testis TAF7L KO model. TAF7L is a germ cell-specific transcription factor. We showed that TAF7L could directly bind to the promoter regions of many CT antigens to regulate their expression in normal testes, and also potentially in cancers. To our surprise, the role of TAF7L resembled that of the previously identified DNA-binding protein BORIS (also known as CTCFL). BORIS per se is a CT antigen whose expression, similar to TAF7L, is also regulated by DNA methylation in its promoter region in different cancer types [78, 79]. Of interest, BORIS can also directly bind to the promoters of other CT antigens, such as NY-ESO-1 in lung cancer cells, to regulate their expression [80]. Thus, there is a subset of CT antigens that function as transcriptional regulators of other CT antigens. Future analysis may identify additional CT antigens with similar roles to provide a more comprehensive outlook on the regulation of CT antigens in cancer.

## Summary

In summary, we have shown that it is possible to extract useful information on the regulation of CT antigens during tumorigenesis by studying gametogenesis. This approach may also provide a unique gateway to probing the functions of these genes in future studies. By taking advantage of the ever-growing publicly available datasets, systematic analysis can, without a doubt, provide new insights into these CT antigens.

**Acknowledgements** This work was supported in part by a grant from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (R01 HD056034 to C.Y.C.); H.C. was supported by the S.Y. Law Memorial Fellowship and the F. Lau Memorial Fellowship

## References

1. Ribas, A., Butterfield, L. H., Glaspy, J. A., & Economou, J. S. (2003). Current developments in cancer vaccines and cellular immunotherapy. *Journal of Clinical Oncology*, *21*, 2415–2432.
2. Rosenberg, S. A., Yang, J. C., & Restifo, N. P. (2004). Cancer immunotherapy: Moving beyond current vaccines. *Nature Medicine*, *10*, 909–915.
3. Tang, E. I., Robinson, C. L., Chong, C. N., Chen, S., & Cheng, C. Y. (2018). A look into the testis as a reservoir for HIV an ZIKV - A reproductive biologist's perspective. In C. Y. Cheng (Ed.), *Spermatogenesis: Biology and clinical implications* (pp. 183–190). CRC Press/Taylor & Francis Group.
4. Meinhardt, A., & Hedger, M. P. (2011). Immunological, paracrine and endocrine aspects of testicular immune privilege. *Molecular and Cellular Endocrinology*, *335*, 60–68.
5. Cheng, Y. H., Wong, E. W., & Cheng, C. Y. (2011). Cancer/testis (CT) antigens, carcinogenesis and spermatogenesis. *Spermatogenesis*, *1*, 209–220.
6. Simpson, A. J., Caballero, O. L., Jungbluth, A., Chen, Y. T., & Old, L. J. (2005). Cancer/testis antigens, gametogenesis and cancer. *Nature Reviews. Cancer*, *5*, 615–625.
7. van der Bruggen, P., Traversari, C., Chomez, P., Lurquin, C., De Plaen, E., Van den Eynde, B., Knuth, A., & Boon, T. (1991). A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. *Science*, *254*, 1643–1647.
8. Almeida, L. G., Sakabe, N. J., deOliveira, A. R., Silva, M. C., Mundstein, A. S., Cohen, T., Chen, Y. T., Chua, R., Gurung, S., Gnjatich, S., Jungbluth, A. A., Caballero, O. L., Bairoch, A., Kiesler, E., White, S. L., Simpson, A. J., Old, L. J., Camargo, A. A., & Vasconcelos, A. T. (2009). CTdatabase: A knowledge-base of high-throughput and curated data on cancer-testis antigens. *Nucleic Acids Research*, *37*, D816–D819.
9. Gjerstorff, M. F., Andersen, M. H., & Ditzel, H. J. (2015). Oncogenic cancer/testis antigens: Prime candidates for immunotherapy. *Oncotarget*, *6*, 15772–15787.
10. Vansteenkiste, J. F., Cho, B. C., Vanakesa, T., De Pas, T., Zielinski, M., Kim, M. S., Jassem, J., Yoshimura, M., Dahabreh, J., Nakayama, H., Havel, L., Kondo, H., Mitsudomi, T., Zarogoulidis, K., Gladkov, O. A., Udud, K., Tada, H., Hoffman, H., Bugge, A., ... Altorki, N. (2016). Efficacy of the MAGE-A3 cancer immunotherapeutic as adjuvant therapy in patients with resected MAGE-A3-positive non-small-cell lung cancer (MAGRIT): A randomised, double-blind, placebo-controlled, phase 3 trial. *The Lancet Oncology*, *17*, 822–835.
11. Ogi, C., & Aruga, A. (2013). Immunological monitoring of anticancer vaccines in clinical trials. *Oncoimmunology*, *2*, e26012.

12. Old, L. J. (2001). Cancer/testis (CT) antigens - A new link between gametogenesis and cancer. *Cancer Immunity*, 1, 1.
13. Wang, C., Gu, Y., Zhang, K., Xie, K., Zhu, M., Dai, N., Jiang, Y., Guo, X., Liu, M., Dai, J., Wu, L., Jin, G., Ma, H., Jiang, T., Yin, R., Xia, Y., Liu, L., Wang, S., Shen, B., ... Hu, Z. (2016). Systematic identification of genes with a cancer-testis expression pattern in 19 cancer types. *Nature Communications*, 7, 10499.
14. Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., & Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protocols*, 11, 1650–1667.
15. Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. *Nature Methods*, 12, 357–360.
16. Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., & Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology*, 33, 290–295.
17. Frazee, A. C., Pertea, G., Jaffe, A. E., Langmead, B., Salzberg, S. L., & Leek, J. T. (2015). Ballgown bridges the gap between transcriptome assembly and expression analysis. *Nature Biotechnology*, 33, 243–246.
18. Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9, 357–359.
19. Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C., Myers, R. M., Brown, M., Li, W., & Liu, X. S. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biology*, 9, R137.
20. Yu, G., Wang, L. G., & He, Q. Y. (2015). ChIPseeker: An R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics*, 31, 2382–2383.
21. Thorvaldsdottir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, 14, 178–192.
22. Robinson, J. T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nature Biotechnology*, 29, 24–26.
23. Whyte, W. A., Orlando, D. A., Hnisz, D., Abraham, B. J., Lin, C. Y., Kagey, M. H., Rahl, P. B., Lee, T. I., & Young, R. A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*, 153, 307–319.
24. Loven, J., Hoke, H. A., Lin, C. Y., Lau, A., Orlando, D. A., Vakoc, C. R., Bradner, J. E., Lee, T. I., & Young, R. A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell*, 153, 320–334.
25. Hart, T., Chandrashekar, M., Aregger, M., Steinhart, Z., Brown, K. R., MacLeod, G., Mis, M., Zimmermann, M., Fradet-Turcotte, A., Sun, S., Mero, P., Dirks, P., Sidhu, S., Roth, F. P., Rissland, O. S., Durocher, D., Angers, S., & Moffat, J. (2015). High-resolution CRISPR screens reveal fitness genes and genotype-specific cancer liabilities. *Cell*, 163, 1515–1526.
26. Perez-Rico, Y. A., Boeva, V., Mallory, A. C., Bitetti, A., Majello, S., Barillot, E., & Shkumatava, A. (2017). Comparative analyses of super-enhancers reveal conserved elements in vertebrate genomes. *Genome Research*, 27, 259–268.
27. Hnisz, D., Abraham, B. J., Lee, T. I., Lau, A., Saint-Andre, V., Sigova, A. A., Hoke, H. A., & Young, R. A. (2013). Super-enhancers in the control of cell identity and disease. *Cell*, 155, 934–947.
28. Kufer, T. A., Sillje, H. H., Korner, R., Gruss, O. J., Meraldi, P., & Nigg, E. A. (2002). Human TPX2 is required for targeting Aurora-A kinase to the spindle. *The Journal of Cell Biology*, 158, 617–623.
29. Gruss, O. J., Wittmann, M., Yokoyama, H., Pepperkok, R., Kufer, T., Sillje, H., Karsenti, E., Mattaj, I. W., & Vernos, I. (2002). Chromosome-induced microtubule assembly mediated by TPX2 is required for spindle formation in HeLa cells. *Nature Cell Biology*, 4, 871–879.

30. Gruss, O. J., Carazo-Salas, R. E., Schatz, C. A., Guarguaglini, G., Kast, J., Wilm, M., Le Bot, N., Vernos, I., Karsenti, E., & Mattaj, I. W. (2001). Ran induces spindle assembly by reversing the inhibitory effect of importin alpha on TPX2 activity. *Cell*, *104*, 83–93.
31. Petry, S., Groen, A. C., Ishihara, K., Mitchison, T. J., & Vale, R. D. (2013). Branching microtubule nucleation in *Xenopus* egg extracts mediated by augmin and TPX2. *Cell*, *152*, 768–777.
32. Guo, F., Yan, L., Guo, H., Li, L., Hu, B., Zhao, Y., Yong, J., Hu, Y., Wang, X., Wei, Y., Wang, W., Li, R., Yan, J., Zhi, X., Zhang, Y., Jin, H., Zhang, W., Hou, Y., Zhu, P., ... Qiao, J. (2015). The transcriptome and DNA methylome landscapes of human primordial germ cells. *Cell*, *161*, 1437–1452.
33. Yamaguchi, S., Hong, K., Liu, R., Shen, L., Inoue, A., Diep, D., Zhang, K., & Zhang, Y. (2013). Tet1 controls meiosis by regulating meiotic gene expression. *Nature*, *492*, 443–447.
34. Yamaguchi, S., Hong, K., Liu, R., Inoue, A., Shen, L., Zhang, K., & Zhang, Y. (2013). Dynamics of 5-methylcytosine and 5-hydroxymethylcytosine during germ cell reprogramming. *Cell Research*, *23*, 329–339.
35. Soumillon, M., Necsulea, A., Weier, M., Brawand, D., Zhang, X., Gu, H., Barthes, P., Kokkinaki, M., Nef, S., Gnirke, A., Dym, M., de Massy, B., Mikkelsen, T. S., & Kaessmann, H. (2013). Cellular source and mechanisms of high transcriptome complexity in the mammalian testis. *Cell Reports*, *3*, 2179–2190.
36. Hunter, A. W., Caplow, M., Coy, D. L., Hancock, W. O., Diez, S., Wordeman, L., & Howard, J. (2003). The kinesin-related protein MCAK is a microtubule depolymerase that forms an ATP-hydrolyzing complex at microtubule ends. *Molecular Cell*, *11*, 445–457.
37. Ganem, N. J., & Compton, D. A. (2004). The KinI kinesin Kif2a is required for bipolar spindle assembly through a functional relationship with MCAK. *The Journal of Cell Biology*, *166*, 473–478.
38. Gnjatich, S., Cao, Y., Reichelt, U., Yekebas, E. F., Nolker, C., Marx, A. H., Erbersdobler, A., Nishikawa, H., Hildebrandt, Y., Bartels, K., Horn, C., Stahl, T., Gout, I., Filonenko, V., Ling, K. L., Cerundolo, V., Luetkens, T., Ritter, G., Friedrichs, K., ... Atanackovic, D. (2010). NY-CO-58/KIF2C is overexpressed in a variety of solid tumors and induces frequent T cell responses in patients with colorectal cancer. *International Journal of Cancer*, *127*, 381–393.
39. Bie, L., Zhao, G., Wang, Y. P., & Zhang, B. (2012). Kinesin family member 2C (KIF2C/MCAK) is a novel marker for prognosis in human gliomas. *Clinical Neurology and Neurosurgery*, *114*, 356–360.
40. Smith, J. S., Tachibana, I., Pohl, U., Lee, H. K., Thanarajasingam, U., Portier, B. P., Ueki, K., Ramaswamy, S., Billings, S. J., Mohrenweiser, H. W., Louis, D. N., & Jenkins, R. B. (2000). A transcript map of the chromosome 19q-arm glioma tumor suppressor region. *Genomics*, *64*, 44–50.
41. Nakamura, Y., Dagio, Y., & Togashi, A. (2009). *Syngn4 for target genes of cancer therapy and diagnosis USA*.
42. Smith, Z. D., & Meissner, A. (2013). DNA methylation: Roles in mammalian development. *Nature Reviews. Genetics*, *14*, 204–220.
43. Schubeler, D. (2015). Function and information content of DNA methylation. *Nature*, *517*, 321–326.
44. Heyn, H., Vidal, E., Ferreira, H. J., Vizoso, M., Sayols, S., Gomez, A., Moran, S., Boque-Sastre, R., Guil, S., Martinez-Cardus, A., Lin, C. Y., Royo, R., Sanchez-Mut, J. V., Martinez, R., Gut, M., Torrents, D., Orozco, M., Gut, I., Young, R. A., & Esteller, M. (2016). Epigenomic analysis detects aberrant super-enhancer DNA methylation in human cancer. *Genome Biology*, *17*, 11.
45. Baubec, T., & Schubeler, D. (2014). Genomic patterns and context specific interpretation of DNA methylation. *Current Opinion in Genetics & Development*, *25*, 85–92.
46. Milekic, M. H., Xin, Y., O'Donnell, A., Kumar, K. K., Bradley-Moore, M., Malaspina, D., Moore, H., Brunner, D., Ge, Y., Edwards, J., Paul, S., Haghghi, F. G., & Gingrich, J. A. (2015). Age-related sperm DNA methylation changes are transmitted to offspring and associated with abnormal behavior and dysregulated gene expression. *Molecular Psychiatry*, *20*, 995–1001.



47. Messerschmidt, D. M., Knowles, B. B., & Solter, D. (2014). DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes & Development*, *28*, 812–828.
48. Seisenberger, S., Andrews, S., Krueger, F., Arand, J., Walter, J., Santos, F., Popp, C., Thienpont, B., Dean, W., & Reik, W. (2012). The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Molecular Cell*, *48*, 849–862.
49. Kagiwada, S., Kurimoto, K., Hirota, T., Yamaji, M., & Saitou, M. (2013). Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. *The EMBO Journal*, *32*, 340–353.
50. Maatouk, D. M., Kellam, L. D., Mann, M. R., Lei, H., Li, E., Bartolomei, M. S., & Resnick, J. L. (2006). DNA methylation is a primary mechanism for silencing postmigratory primordial germ cell genes in both germ cell and somatic cell lineages. *Development*, *133*, 3411–3418.
51. Hargan-Calvopina, J., Taylor, S., Cook, H., Hu, Z., Lee, S. A., Yen, M. R., Chiang, Y. S., Chen, P. Y., & Clark, A. T. (2016). Stage-Specific Demethylation in Primordial Germ Cells Safeguards against Precocious Differentiation. *Developmental Cell*, *39*, 75–86.
52. Li, W., Guo, X. J., Teng, F., Hou, X. J., Lv, Z., Zhou, S. Y., Bi, Y., Wan, H. F., Feng, C. J., Yuan, Y., Zhao, X. Y., Wang, L., Sha, J. H., & Zhou, Q. (2013). Tex101 is essential for male fertility by affecting sperm migration into the oviduct in mice. *Journal of Molecular Cell Biology*, *5*, 345–347.
53. Fujihara, Y., Tokuhira, K., Muro, Y., Kondoh, G., Araki, Y., Ikawa, M., & Okabe, M. (2013). Expression of TEX101, regulated by ACE, is essential for the production of fertile mouse spermatozoa. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 8111–8116.
54. Ghafouri-Fard, S., Abbasi, A., Moslehi, H., Faramarzi, N., Taba Taba Vakili, S., Mobasheri, M. B., & Modarressi, M. H. (2009). Elevated expression levels of testis-specific genes TEX101 and SPATA19 in basal cell carcinoma and their correlation with clinical and pathological features. *The British Journal of Dermatology*, *162*, 772–779.
55. Dianatpour, M., Mehdipour, P., Nayernia, K., Mobasheri, M. B., Ghafouri-Fard, S., Savad, S., & Modarressi, M. H. (2012). Expression of testis specific genes TSGA10, TEX101 and ODF3 in breast cancer. *Iranian Red Crescent Medical Journal*, *14*, 722–726.
56. Mobasheri, M. B., Shirkoobi, R., Zendehdel, K., Jahanzad, I., Talebi, S., Afsharpad, M., & Modarressi, M. H. (2015). Transcriptome analysis of the cancer/testis genes, DAZ1, AURKC, and TEX101, in breast tumors and six breast cancer cell lines. *Tumour Biology*, *36*, 8201–8206.
57. Yoshitake, H., Yokoi, H., Ishikawa, H., Maruyama, M., Endo, S., Nojima, M., Yoshida, K., Yoshikawa, H., Suzuki, F., Takamori, K., Fujiwara, H., & Araki, Y. (2012). Overexpression of TEX101, a potential novel cancer marker, in head and neck squamous cell carcinoma. *Cancer Biomarkers*, *12*, 141–148.
58. Heiser, L. M., Sadanandam, A., Kuo, W. L., Benz, S. C., Goldstein, T. C., Ng, S., Gibb, W. J., Wang, N. J., Ziyad, S., Tong, F., Bayani, N., Hu, Z., Billig, J. I., Dueregger, A., Lewis, S., Jakkula, L., Korkola, J. E., Durinck, S., Pepin, F., ... Spellman, P. T. (2012). Subtype and pathway specific responses to anticancer compounds in breast cancer. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 2724–2729.
59. Pointud, J. C., Mengus, G., Brancorsini, S., Monaco, L., Parvinen, M., Sassone-Corsi, P., & Davidson, I. (2003). The intracellular localisation of TAF7L, a paralogue of transcription factor TFIID subunit TAF7, is developmentally regulated during male germ-cell differentiation. *Journal of Cell Science*, *116*, 1847–1858.
60. Kimmins, S., Kotaja, N., Davidson, I., & Sassone-Corsi, P. (2004). Testis-specific transcription mechanisms promoting male germ-cell differentiation. *Reproduction*, *128*, 5–12.
61. Cheng, Y., Buffone, M. G., Kouadio, M., Goodheart, M., Page, D. C., Gerton, G. L., Davidson, I., & Wang, P. J. (2007). Abnormal sperm in mice lacking the Taf7l gene. *Molecular and Cellular Biology*, *27*, 2582–2589.

62. Yazarloo, F., Shirkoohi, R., Mobasheri, M. B., Emami, A., & Modarressi, M. H. (2013). Expression analysis of four testis-specific genes AURKC, OIP5, PIWIL2 and TAF7L in acute myeloid leukemia: A gender-dependent expression pattern. *Medical Oncology*, *30*, 368.
63. Zhou, H., Grubisic, I., Zheng, K., He, Y., Wang, P. J., Kaplan, T., & Tjian, R. (2013). Taf7l cooperates with Trf2 to regulate spermiogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 16886–16891.
64. Dong, X. Y., Su, Y. R., Qian, X. P., Yang, X. A., Pang, X. W., Wu, H. Y., & Chen, W. F. (2003). Identification of two novel CT antigens and their capacity to elicit antibody response in hepatocellular carcinoma patients. *British Journal of Cancer*, *89*, 291–297.
65. Doghman-Bouguerra, M., Granatiero, V., Sbiera, S., Sbiera, I., Lacas-Gervais, S., Brau, F., Fassnacht, M., Rizzuto, R., & Lalli, E. (2016). FATE1 antagonizes calcium- and drug-induced apoptosis by uncoupling ER and mitochondria. *EMBO Reports*, *17*, 1264–1280.
66. da Silva, V. L., Fonseca, A. F., Fonseca, M., da Silva, T. E., Coelho, A. C., Kroll, J. E., de Souza, J. E. S., Stransky, B., de Souza, G. A., & de Souza, S. J. (2017). Genome-wide identification of cancer/testis genes and their association with prognosis in a pan-cancer analysis. *Oncotarget*, *8*, 92966–92977.
67. Old, L. J. (2007). Cancer is a somatic cell pregnancy. *Cancer Immunity*, *7*, 19.
68. Mansour, M. R., Abraham, B. J., Anders, L., Berezovskaya, A., Gutierrez, A., Durbin, A. D., Etchin, J., Lawton, L., Sallan, S. E., Silverman, L. B., Loh, M. L., Hunger, S. P., Sanda, T., Young, R. A., & Look, A. T. (2014). Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. *Science*, *346*, 1373–1377.
69. Chapuy, B., McKeown, M. R., Lin, C. Y., Monti, S., Roemer, M. G., Qi, J., Rahl, P. B., Sun, H. H., Yeda, K. T., Doench, J. G., Reichert, E., Kung, A. L., Rodig, S. J., Young, R. A., Shipp, M. A., & Bradner, J. E. (2013). Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. *Cancer Cell*, *24*, 777–790.
70. Chipumuro, E., Marco, E., Christensen, C. L., Kwiatkowski, N., Zhang, T., Hatheway, C. M., Abraham, B. J., Sharma, B., Yeung, C., Altabef, A., Perez-Atayde, A., Wong, K. K., Yuan, G. C., Gray, N. S., Young, R. A., & George, R. E. (2014). CDK7 inhibition suppresses super-enhancer-linked oncogenic transcription in MYCN-driven cancer. *Cell*, *159*, 1126–1139.
71. Kulis, M., & Esteller, M. (2010). DNA methylation and cancer. *Advances in Genetics*, *70*, 27–56.
72. Van Tongelen, A., Lorient, A., & De Smet, C. (2017). Oncogenic roles of DNA hypomethylation through the activation of cancer-germline genes. *Cancer Letters*, *396*, 130–137.
73. De Smet, C., & Lorient, A. (2013). DNA hypomethylation and activation of germline-specific genes in cancer. *Advances in Experimental Medicine and Biology*, *754*, 149–166.
74. Shichijio, S., Yamada, A., Sagawa, K., Iwanamoto, O., Sakata, M., Nagai, K., & Itoh, K. (1996). Induction of MAGE genes in lymphoid cells by the demethylating agent 5-aza-2'-deoxycytidine. *Japanese Journal of Cancer Research*, *87*, 751–756.
75. Milutinovic, S., Zhuang, Q., Niveleau, A., & Szyf, M. (2003). Epigenomic stress response. Knockdown of DNA methyltransferase 1 triggers an intra-S-phase arrest of DNA replication and induction of stress response genes. *The Journal of Biological Chemistry*, *278*, 14985–14995.
76. James, S. R., Link, P. A., & Karpf, A. R. (2006). Epigenetic regulation of X-linked cancer/germline antigen genes by DNMT1 and DNMT3b. *Oncogene*, *25*, 6975–6985.
77. De Smet, C., Lurquin, C., Lethe, B., Martelange, V., & Boon, T. (1999). DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter. *Molecular and Cellular Biology*, *19*, 7327–7335.
78. Hoffmann, M. J., Muller, M., Engers, R., & Schulz, W. A. (2006). Epigenetic control of CTCFL/BORIS and OCT4 expression in urogenital malignancies. *Biochemical Pharmacology*, *72*, 1577–1588.



79. Woloszynska-Read, A., James, S. R., Link, P. A., Yu, J., Odunsi, K., & Karpf, A. R. (2007). DNA methylation-dependent regulation of BORIS/CTCF expression in ovarian cancer. *Cancer Immunity*, 7, 21.
80. Hong, J. A., Kang, Y., Abdullaev, Z., Flanagan, P. T., Pack, S. D., Fischette, M. R., Adnani, M. T., Loukinov, D. I., Vatolin, S., Risinger, J. I., Custer, M., Chen, G. A., Zhao, M., Nguyen, D. M., Barrett, J. C., Lobanekov, V. V., & Schrump, D. S. (2005). Reciprocal binding of CTCF and BORIS to the NY-ESO-1 promoter coincides with derepression of this cancer-testis gene in lung cancer cells. *Cancer Research*, 65, 7763–7774.