HAL open science

Combined RNA/tissue profiling identifies novel Cancer/Testis genes

Soazik Jamin, Feria Hikmet, Romain Mathieu, Bernard Jégou, Cecilia Lindskog, Frédéric Chalmel, Michael Primig

To cite this version:

Soazik Jamin, Feria Hikmet, Romain Mathieu, Bernard Jégou, Cecilia Lindskog, et al.. Combined RNA/tissue profiling identifies novel Cancer/Testis genes. Molecular Oncology, 2021, 15 (11), pp.3003-3023. 10.1002/1878-0261.12900 . hal-03163968

HAL Id: hal-03163968

https://hal.science/hal-03163968

Submitted on 8 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Check for updates

Accepted Article

DR. MICHAEL PRIMIG (Orcid ID : 0000-0002-2061-0119)

Received Date : 20-Jul-2020
Revised Date : 19-Nov-2020
Accepted Date : 24-Dec-2020
Article type : Research Article

Combined RNA/tissue profiling identifies novel Cancer/Testis genes

Soazik P. Jamin[1], Feria Hikmet[2], Romain Mathieu[1,3], Bernard Jégou[1], Cecilia Lindskog[2], Frédéric Chalmel[1,#] and Michael Primig[1,#,*]

[1]Univ Rennes, Inserm, EHESP, Irset (Institut de recherche en santé, environnement et travail) – UMR_S 1085, F-35000 Rennes, France

[2]Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Rudbeck Laboratory, Uppsala University, 75185 Uppsala, Sweden

[3]Department of Urology, University Hospital, 35000 Rennes, France

[#]These authors made an equal contribution

[*]Corresponding author: Michael Primig; IRSET; 9, rue du Prof. Léon Bernard, 35000 Rennes, France. Phone: +33 2 23 23 61 78 | Email: michael.primig@inserm.fr

Running title: Novel Cancer/Testis genes with oncogenic potential

Keywords: Cancer/Testis genes, oncogenes, GeneChips, RNA-Seq, Tissue Microarrays

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/1878-0261.12900

Molecular Oncology (2020) © 2020 The Authors. Published by FEBS Press and John Wiley & Sons Ltd.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Abbreviations:

Cancer/Testis genes: CT genes

Tissue Microarrays: TMAs

Accepted Article

Molecular Oncology (2020) © 2020 The Authors. Published by FEBS Press and John Wiley & Sons Ltd.

Abstract

Cancer/Testis (CT) genes are induced in germ cells, repressed in somatic cells and de-repressed in somatic tumors, where these genes can contribute to cancer progression. CT gene identification requires data obtained using standardized protocols and technologies. This is a challenge because data for germ cells, gonads, normal somatic tissues and a wide range of cancer samples stem from multiple sources and were generated over substantial periods of time.

We carried out a GeneChip-based RNA profiling analysis using our own data for testis and enriched germ cells, data for somatic cancers from the Expression Project for Oncology and data for normal somatic tissues from the Gene Omnibus repository. We identified 478 candidate loci that include known CT genes, numerous genes associated with oncogenic processes and novel candidates that are not referenced in the Cancer/Testis Database (www.cta.lncc.br). We complemented RNA expression data at the protein level for SPESP1, GALNTL5, PDCL2 and C11orf42 using cancer tissue microarrays covering malignant tumors of breast, uterus, thyroid and kidney, as well as published RNA profiling and immunohistochemical data provided by the Human Protein Atlas (www.proteinatlas.org).

We report that combined RNA/tissue profiling identifies novel CT genes that may be of clinical interest as therapeutical targets or biomarkers. Our findings also highlight the challenges of detecting truly germ cell-specific mRNAs and the proteins they encode in highly heterogenous testicular, somatic and tumor tissues.

Accepted Article

Molecular Oncology (2020) © 2020 The Authors. Published by FEBS Press and John Wiley & Sons Ltd.

1. Introduction

Cancer/Testis (CT) genes are expressed in testicular cells and in somatic cancers but typically not in their corresponding normal somatic tissues; reviewed in [1]. The majority of CT genes function in gametogenesis and fertility and their abnormal expression in somatic cancer cells can contribute to malignant properties. Indeed, previous work has identified CT genes that are essential for cancer cell division and that affect regulatory signal transduction pathways [2]. Gaining insight into CT gene's biological roles helps better understand how cancer cells proliferate, form metastases, repair DNA damage, suppress apoptosis, alter signaling pathways and invade normal tissues; for review, see [3]. More recently, CT genes were proposed to be biomarkers for cancer stem cells that are thought to play roles in the maintenance of tumor growth and resistance to chemotherapy; reviewed in [4].

Molecular biological and genomic approaches have led to the discovery of several hundred CT genes referenced in the CTdatabase (www.cta.lncc.br) [5]. These loci were initially classified into testis-specific, testis/brain-specific and testis-selective groups (whereby low expression in two non-testicular tissues is observed); it is noteworthy that the majority of CT genes has been identified via mRNA expression in limited somatic control sample sets, which may explain why they are frequently not genuinely testis-specific; for a detailed discussion, see [6]. It is therefore currently unclear how many CT genes are indeed expressed only in male gonads and somatic cancers both at the RNA and protein levels. More recent work carried out by the group developing the Human Protein Atlas, that monitors protein localization in all normal human tissues and numerous cancers, has made a major contribution to determining the human testicular proteome. This study also revealed the somatic cancer expression profile of testicular proteins, some of which are relevant for non-small-cell lung cancer [7, 8].

We report a combined microarray-based RNA/tissue profiling analysis of somatic cancers, normal tissues, prepubertal and adult testis biopsies, total testis samples and enriched meiotic and post-meiotic germ cells. The approach identified most known and novel CT genes in addition to oncogenes and cancer-associated genes that had not been profiled in the male germline before. We selected promising cases and further characterized the proteins they encode using testicular sections and cancer tissue microarrays. Our results represent a rich source for further functional analyses of CT genes in the field of molecular oncology.

Accepted Article

2. Materials and Methods

GeneChip RNA profiling data assembly

The entire dataset was generated with Affymetrix Human Genome U133 Plus 2.0 GeneChips (ThermoFisher). Expression data for human testis (total testis and isolated seminiferous tubules), prepubertal biopsies (high-, intermediate- and low infertility risk; HIR, IIR and LIR), adult testicular biopsies with different Johnson Scores indicating the steps where spermatogenesis is disrupted (JS1, 2, 3, 5, 7, 8 and 10) and enriched germ cells (pachytene spermatocytes and round spermatids) were described in reference [9]. Expression data for 45 normal somatic control tissues were downloaded from the NCBI's Gene Omnibus (GEO; GSE7307, GSE6565, GSE11839) repository [10]; see Supplemental File S1 columns BZ to DR for tissue type annotation data. Cancer expression data for 214 cancer sub-types produced by the Expression for Oncology (expO) project (www.intgen.org) were retrieved from GEO (GSE2109) and combined with two other datasets (GSE10802 and GSE6891); see Supplemental File S1 columns DT to LN for cancer types.

GeneChip data processing and analysis

GeneChip U133 Plus 2.0 expression data were quality controlled, processed and normalized as in reference [9, 11]. After quality control, raw data CEL files were normalized, background corrected and summarized with the Robust MultiArray Average (RMA) function implemented in AMEN [12].

Next, we defined classes of transcripts according to their expression pattern in normal and cancer tissue samples. Transcripts specifically expressed (SE) or preferentially expressed (PE) in a given normal tissue were identified by applying three filtration steps. First, intensity signal is above the background expression cutoff (BEC = 5.5, corresponding to the overall median log2-transformed intensity) in the tissue of interest and below this threshold in all the other normal tissues (with three exception for PE transcripts). Second, we required at least a 2-fold change between the signal in the tissue of interest and those of all other tissues with three exceptions for PE transcripts. Third, statistically significant changes across the samples were identified using a LIMMA statistical test with the false discovery rate (FDR) adjustment method ($p \leq 0.01$).

Accepted Article

Transcripts preferentially or specifically expressed in testis as compared to the other somatic normal tissues are termed PET and SET, respectively. SEHET (specifically expressed and highly expressed in testis) and PEHET (preferentially expressed and highly expressed in testis) transcripts show expression signal in testis or germ cells that are in the upper quartile of the overall log2-transformed expression matrix.

Transcripts upregulated in a given cancer subtype (UC) were also identified by applying three filtration steps. First, the intensity signal was above the background expression cutoff (BEC = 5.5, corresponding to the overall median log2-transformed intensity) in the cancer subtype of interest. Second, a 2-fold change between the signal in the cancer subtype and that of the corresponding somatic tissue. Third, the statistically significant changes across the samples were identified using a LIMMA statistical test with the false discovery rate (FDR) adjustment method ($p \leq 0.01$). UHEC (upregulated and highly expressed in cancer) correspond to UC transcripts but for which the expression signal in a given cancer subtype is in the upper quartile. UCNDH (upregulated in cancer and not detected in healthy tissues) transcripts correspond to UC transcripts but for which the expression signal in the corresponding somatic tissue is below the BEC. Finally, UHECNDH (upregulated and highly expressed in cancer and not detected in healthy tissues) transcripts correspond to UC transcripts but for which the expression signal in a given cancer subtype is in the upper quartile and the expression signal in the corresponding somatic tissue is below the BEC.

Bulk and single-cell RNA-Sequencing data processing and visualization

We integrated RNA-sequencing data for total testis samples and enriched testicular cells (Sertoli cells, Leydig cells, peritubular cells, spermatocytes and round spermatids) published by [13, 14] and single-cell RNA sequencing (scRNA-Seq) data for adult testicular cells published by [15]. The single-cell expression profiles (scatter plots) for individual genes were generated online *via* the Reproductive Genomics Viewer (https://rgv.genouest.org) [16, 17].

RNA-Sequencing data for total testis, somatic tissues and cancer samples from the Human Proteome Atlas were processed by applying pseudo-counts of +1 and log2 transformation to the dataset (www.proteinatlas.org; [18]). The signals were visualized using the *heatmap.2* package in R (CRAN); samples and genes were grouped using the default *hclust* algorithm using default color

Accepted Article

scaling for normal and testis tissues and row-wise color scaling for cancer data provided by the TCGA consortium [19].

Immunohistochemistry analysis using testicular sections

Regarding human samples, the local ethics committee approved the experimental protocol ‘‘Study of Normal and Pathological Human Spermatogenesis’’ (registration PFS09-015) at the French Biomedicine Agency; informed consent was obtained from all donors. The study’s methodologies adhere to the standards set by the Declaration of Helsinki.

Immunohistochemical analyses using human testicular sections were carried out using a standard protocol as described [20]. Briefly, sections were deparaffinized in Ottix Plus (MM France Microm) and then rehydrated in ethanol and distilled water. Antigen retrieval was performed in citrate buffer pH 6.0 (Fisher Scientific) for 20 minutes at 95°C. The tissues were saturated in Ultra Vision Block (Fisher Scientific) for 5 minutes and antibodies against SPESP1 (1:5000; Sigma HPA 051040), GALNTL5 (1:500; Sigma HPA 011140), PDCL2 (1:2500; Sigma HPA 048260) and C11orf42 (1:175; Sigma HPA 063404) were applied for 16h at 4°C in a humidified chamber. The samples were washed and exposed to Primary Antibody Enhancer (Fisher Scientific) for 10 minutes and HRP polymer (Fisher Scientific) for 15 minutes. The substrate DAB quanto (Fisher Scientific) was applied and the sections were counterstained with hematoxylin (Leica). The slides were scanned at the H2P2 platform (Biosit, Rennes).

Immunohistochemical protein detection using Tissue MicroArrays

The IHC protocol that we used to analyse testicular sections was also applied to TMAs (Biomax, US) that were employed to analyse SPESP1 (SK803a) GALNTL5 (EM1021, UT721, HThy-Pap120CS-01), PDCL2 (KD1503) and C11orf42 (TH481).

Accepted Article

3. Results

3.1. Experimental rationale

In earlier work we used GeneChips to determine the testicular transcriptome in biopsies from pre-pubertal and adult individuals, total testicular samples and adult meiotic and post-meiotic germ cells. These analyses included an extensive set of somatic samples to classify genes into testis-specific, preferentially expressed in testis and ubiquitous [9, 21]. We integrated our data with normal somatic controls from the NCBI's GEO repository and high-quality GeneChip cancer expression data provided by the Expression for Oncology (expO) project (www.intgen.org) [22]. Our RNA profiling study is thus based on robust expression data from multiple sources that were produced using a highly standardized RNA profiling method. Our work also distinguishes itself from similar analyses by comprehensive total testis, testicular biopsy and male germ cell sampling in combination with antibody-based tissue profiling [23-25]. We analysed 61 testis-associated samples including total testis (2 samples), seminiferous tubules (2), meiotic (2) and post-meiotic (2) germ cells, pubertal (15) and adult (38) testicular biopsies. Furthermore, we processed 544 samples from 45 normal somatic tissues that we used as controls and 2281 samples corresponding to 214 cancer subtypes from 23 distinct tissue origins.

3.2. CT gene identification and definition of expression level cutoff values

We assembled our testicular and germ cell data [9], cancer data from the expO project and data for normal somatic control tissues from the GEO repository (Figure 1A). The dataset comprises 2998 GeneChips (Human Genome U133 Plus 2.0) among which 2887 passed quality control (see methods for details; Figure 1A). They were processed, normalized and used in a differential gene expression analysis like in reference [9]. The median log2 expression value was set at 5.5 and lower and upper boundaries between the 25th and 75th percentile defining a window of gene expression are 4.4 and 6.9, respectively. Values above the 75th percentile represent high expression and transcripts associated with values below the 25th percentile were considered to be undetectable (Figure 1B). In step 1, among 54613 probesets we selected 2140 (corresponding to 1433 genes) that displayed significant expression in testis or germ cells, including 1285 probesets that were specifically expressed in testis. In step 2, we identified 2819 probesets (2025 genes) as being upregulated in at least one somatic cancer subtype and not expressed in the corresponding

Accepted Article

normal somatic tissue (UCNDH). The intersection of steps 1 and 2 identified 602 probe sets (478 genes) that displayed a pattern broadly corresponding to CT genes (Figure 2; see filtering options in Supplemental File S1).

3.3. Extended somatic control sample sets are critical for identifying bona fide CT genes

Many putative CT genes are not testis-specific [6]. We therefore determined expression patterns of 176 probeset-associated mRNAs referenced in the CTDatabase using our sample set (www.cta.lncc.br [5]). We find that 39% (69/176 probesets) are specifically expressed and highly expressed in testis (SEHET) and 12% (22/176) are specifically expressed in testis (SET) (Figure 3; CT Genes referenced in CTDatabase are available via Supplemental File S1). However, 14% (25/176) are only preferentially expressed and highly expressed in testis (PEHET) and 5% (8/176) are preferentially expressed (PET), which indicates that their mRNA is detected in at least one somatic sample. For 13% (22/176) we find intermediate expression (IE, expressed in 4-10 somatic tissues) and a large group of 16% (28/176) even show ubiquitous expression (UE) in all somatic controls (Figure 3). This pattern is unsurprising because testis contain not only germ cells but also Sertoli cells, Leydig cells, peritubular cells, smooth muscle cells and immune cells. We note that among 1079 testicular proteins, 261 were also detected in 22 normal somatic tissues, including fallopian tube (109), cerebral cortex (46 proteins) and epididymis (28) (Figure S1) [7]. These results underline that comprehensive control sample sets are critical for the identification of testis-specific CT genes.

3.4. RNA profiling using GeneChip and RNA-Sequencing data identifies novel CT genes

To validate our filtration method we selected 182 probesets (corresponding to 124 unique genes) that show testis-specific expression (SET and SEHET classes), up-regulation in at least one cancer subtype and no expression in the corresponding somatic tissue (UCNDH) in our dataset (select SET plus SEHET and UCNDH filter options in columns G and L, respectively, in Supplemental File S1). To confirm and extend our initial GeneChip expression data, we analysed testicular expression with our RNA-Sequencing data from total testis samples and enriched meiotic spermatocytes, post-meiotic round spermatids as well as Sertoli-, Leydig- and peritubular cells that were available for 115 core CT genes [13, 14]. As expected, we found that the vast majority

Accepted Article

of the core CT genes are highly induced in the male germline (see RNA-Seq data in Figure S2A). We then extended the analysis using single-cell RNA-Sequencing data for testicular somatic cells (Sertoli-, Leydig- and peritubular cells and macrophages), mitotic germ cells (dividing, differentiating and differentiated spermatogonia), meiotic germ cells (leptotene, zygotene, pachytene, diplotene and diakinesis spermatocytes) and post-meiotic germ cells (spermatids) [15]. The result confirmed that nearly all CT genes are expressed in germ cells at different stages of differentiation, including mitotic, meiotic and post-meiotic phases of male gametogenesis (see scRNA-Seq data in Figure S2B).

Next, we confirmed the testis-specific or testis-enriched expression pattern determined with GeneChip data for core CT genes by using RNA-Sequencing data available to us [18]. We compared expression levels in male and female gonads to 35 normal somatic tissues and found that the majority of the genes shows the expected testis-specific, preferential or testis/brain expression patterns (Figure 4A; see Supplemental File S2 for gene annotation and signal intensities). This result underlines that GeneChip RNA profiling data obtained with samples that were processed using highly standardized methods are reproducible in the majority of the cases, even across RNA profiling methods based on fundamentally different technologies. It is unclear why some transcripts detectable in normal somatic tissues by RNA-Sequencing fail to be scored as expressed by GeneChips. Different threshold levels of detection, signal processing procedures and evolving genome annotation data that analysis procedures are based on may be at least in part explain the discrepancies.

Finally, we analysed RNA-Sequencing data from the TCGA consortium to explore the expression patterns of core CT genes in testicular and ovarian cancer versus 15 selected somatic malignancies [19, 26]. We again found the majority of them to be expressed in at least one tumor. The identification of CT genes in ovarian and testicular cancer points to meiotic functions shared by male and female gonads (Figure 4B).

The core CT genes include CMTM1 [27], CT83 (CXorf61) [28], EZHIP (CXorf67) [29], DCAF4L2 [30], KHDRBS3 [31], LEMD1 [32], PIWIL1 [33] and SPATA6 [34], which contribute to cancer progression and metastasis. We also find HORMAD1 [35], TDRD6 [36], ZFAND4 (ANUBL1) [37], SOHLH2 [38], STK31 [39], TEX9 [40] and USP44 [41], which affect cell growth or are prognostic biomarkers. This indicates that the output of our approach is biomedically relevant.

Accepted Article

We further investigated expression patterns of novel CT gene candidates, for which the available literature either shows testicular roles or reports critical molecular functions in somatic tumors but not both. The SEHET class gene DCAF4L1 (DDB1- and CUL4-associated factor 4-like protein 1) has no currently annotated molecular function. However, its mRNA is testis-specific in our sample set and peaks in embryonic ovary germ cell tumor and adult seminomas (compare RNA-Sequencing data in Figure 4 with GeneChip data in Figure 5A). Interestingly, genetic variations in this locus were associated with hemangioblastoma, a benign brain tumor [42]. DCAF4L1 shows significant differential expression in 14 cancers versus normal controls and is detected in ovarian and testicular germ cell tumors. This includes kidney cancer and bladder cancer. Interestingly, high expression of DCAF4L1 is associated with a decreased probability for survival in kidney cancer patients and an increased probability in the case of bladder cancer (TCGA consortium, http://timer.cistrome.org; [43], Figures S3A and S4A).

Another SEHET-type gene is DMRT1 (Doublesex- and mab-3-related transcription factor 1), which belongs to a highly conserved family of DNA binding transcription factors important for development and sex differentiation. The mouse protein controls germ stem cell differentiation and the transition from mitotic growth to meiotic development in the germline (for review, see [44]). We find elevated levels of the human mRNA notably in endometrium, ovary and breast cancer samples as well as testicular cancer (Figure 4 and Figure 5B). The latter is in keeping with genetic data that associate DMRT1 with testicular germ cell tumor susceptibility [45, 46]. Immunohistochemical data from the Human Protein Atlas (HPA) confirm that pattern in the case of breast cancer (see www.proteinatlas.org [47]). DMRT1 is up-regulated in nine cancers and also shows strong signals in testicular germ cell cancers in both GeneChip and RNA-Seq data sets. This includes endometrial cancers, for which we detect a strong expression peak that corresponds to the findings reported by the TCGA consortium (Figure 5B, Figure S3B). We note that high expression in Uterine Corpus Endometrial Carcinoma is associated with decreased survival (Figure S4B).

TMEM217 (Transmembrane protein 217) also shows a SEHET pattern, encodes a predicted transmembrane protein and transcriptionally responds to an antiproliferative agent [48]. The mRNA accumulates in leukemia samples in our dataset and in various somatic malignancies, including thyroid cancer, as reported by HPA (Figure 4 and Figure 5C; www.proteinatlas.org; [47]). TMEM217's function and its expression in normal and cancer tissues are currently

Accepted Article

unknown. The gene is significantly differentially expressed in 14 cancers versus controls, including kidney renal papillary cell carcinoma (KIRP) and thyroid carcinoma (THCA) where it is induced. We note that TMEM217 is highly induced in leukemia samples assayed with GeneChips and RNA-Sequencing (Figure 5C, Figure S3C) and patients diagnosed with acute myeloid leukemia (LAML) show a decreased survival rate when the gene is highly expressed (Figure S4C).

The SET-type NLRP7 (NACHT, LRR and PYD domains-containing protein 7) is involved in the inflammatory response and was associated with myometrial invasion in human endometrial cancer [49]. The mRNA peaks in different ovarian tumors and appears to be strongly induced in testicular cancer (Figure 4 and Figure 5D). NLRP7 is differentially expressed in 14 tumors and is also induced in ovarian cancer and testicular germ cell cancers in particular, like in our RNA profiling data sets (Figure 5D and Figure S3D).

Finally, the SEHET-class gene SYCP2 (Synaptonemal complex protein 2) shows a particularly striking CT gene pattern because it is de-repressed in a variety of somatic tumors, especially breast-, and cervical cancer to an unusually high level (Figure 4 and Figure 5E). The protein interacts with other components of the synaptonemal complex, which ensures the separation of homologous chromosomes during the first meiotic division in male germ cells [50]. Normal expression of SYCP2 is essential for male fertility [51]. Importantly, SYCP2 was recently reported to be a biomarker for luminal A/B breast cancer [52]. SYCP2 is differentially expressed in 11 cancers and shows much stronger RNA-Seq signals in breast and cervical cancer as compared to normal controls, which is coherent with the GeneChip data (Figure 5E and Figure S3E).

These five cases exemplify potential CT genes that are promising candidates for functional analyses since they have been broadly associated with cell growth, differentiation and cancer.

3.6. CT gene analysis at the protein level by Tissue MicroArrays

We next sought to further investigate the RNA/protein profiles of new CT genes for which no direct evidence was reported in the scientific literature (referenced in PubMed) that links them to altered (benign or malign) mitotic cell division [22]. To this end, we selected four candidates that showed promising mRNA/protein profiling patterns using our GeneChip expression data and IHC assays from HPA (www.proteinatlas.org [47]). We first employed published scRNA-Sequencing data to explore their expression patterns within testicular tissue [15] (Figure 6A). The results are

Accepted Article

coherent with broad expression in the germline (SPESP1), induction in spermatocytes and peak expression in spermatids (the core gene GALNTL5 and PDCL2) and mostly spermatid-specific expression (C11orf42) (Figure 6B-E).

SPESP1 is an interesting gene because we classified it as specifically and highly expressed in testis (SEHET) and it encodes a membrane protein, which localizes to the sperm acrosome (see Figure 6B for scRNA-Sequencing data and Figure 7A for GeneChip data). Infertile male patients were found to make antibodies against SPESP1 and the mouse ortholog was shown to be involved in male fertility [53-55]. These results suggest a similarly important function in mammalian spermiogenesis and male fertility for human SPESP1. The gene is transcribed in, among several malignancies, skin, liver and vulval cancer, which is confirmed at the protein level for skin and liver cancer by HPA (www.proteinatlas.org; Figure 7A). We first performed an IHC assay using an HPA antibody and confirmed the protein's presence in round and elongated spermatids, which corresponds to the gene's expression in testicular samples and enriched germ cells (Figure 7B); see also (www.proteinatlas.org; [47]). Next, we employed a commercial tissue microarray (TMA) covering 84 benign and malign skin tumors, 12 samples from other tumors (breast, ear, fibrous tissue, parotid gland, vulva) and four normal skin samples to analyse SPESP1's staining pattern (SK803a) (Supplemental Figure S5A). We observed that 67% (62/93) of the cancer samples on a TMA showed variable immunohistochemical signals, while the remaining cases were not stained (Figure 7C and samples A7/8 and C7/8 in panel D). Unexpectedly, we also detected cytoplasmic staining of keratinocytes in four normal skin samples (Figure 7D samples J7/8). The faint signal is also present in epidermal cells of normal skin samples published by HPA that were analysed with the antibody we employed (HPA051040); however, a second antibody (HPA045936) does not yield a signal. This is in contrast to the acrosomal staining pattern that is similar for both antibodies (www.proteinatlas.org; [47]). Our GeneChip data and published RNA-Sequencing data using skin samples do not indicate expression values for SPESP1 that are above background (www.proteinatlas.org; [47]). Moreover, a mouse model lacking *Spesp1* shows male infertility but no defect in any somatic tissues, including skin, which argues in favour of testis-specific roles (hence expression) [54]. It is therefore unclear if weak signals in keratinocytes annotated as normal are true and physiologically relevant. Taken together, the data support the notion that SPESP1 is a testicular protein that accumulates in a substantial fraction of skin cancer samples.

Accepted Article

We next analysed GALNTL5, which encodes a membranous inactive polypeptide N-acetylgalactosaminyltransferase-like protein likely important for acrosome function and proteolysis in sperm [56]. A mutation in the gene was associated with abnormal spermatogenesis in human [57]. Our GeneChip data indicate strong expression of GALNTL5 in adult testis, enriched spermatocytes and round spermatids (Figure 6C and Figure 8A). To test the commercially available HPA antibody (HPA011140), we first assayed GALNTL5 on adult testicular sections and observed membrane and cytoplasmic staining in spermatocytes, nuclear staining in round spermatids and cytoplasmic staining in Leydig cells (Figure 8B). These signals broadly correspond to patterns displayed in HPA (www.proteinatlas.org; [47]). GALNTL5's RNA/protein expression profile is thus consistent with a role in the male germline as suggested by earlier work [57, 56].

GeneChip profiling data for GALNTL5 show a classical CT gene pattern in an adenocarcinoma of the endometrium. HPA's data confirm expression in endometrium at the protein level and also reveal thyroid and skin cancer samples as positively stained by IHC (Figure 8A; www.proteinatlas.org; [47]). We therefore confirmed the testicular expression of GALNTL5 using sections (Figure 8B) and then we employed tumor TMAs for endometrium (EM1021a, UT721) and thyroid papillary carcinoma (HThy-Pap120CS-01) for further analysis (Supplemental Figure S5B). Consistently, we detected cytoplasmic staining of variable intensity in 23/97 (24%) endometrium cancer samples. We also detected weak staining in 1/5 (20%) controls (EM1021a, Figure 8C cancer samples G8-11 and normal controls H8-11). In a similar experiment using a different TMA 33/83 (40%) malignant samples contained detectable levels of the protein while 0/9 controls were stained (UT721 A1, Figure 8D cancer samples E and F1-4). We also observed that 5/8 (63%) of the normal adjacent tissue (NAT) controls showed variable levels of staining. This again suggests that histologically normal tissues can transcribe and translate CT genes and therefore possess a molecular feature that may mark them out as (pre)-malignant in spite of their normal histological appearance. Finally, we detected typically strong cytoplasmic staining in 54/58 (93%) cancerous thyroid samples, while only 14/62 (23%) of the NAT samples showed almost exclusively weak signals (HThy-Pap120CS-01; Figure 8E cancer samples A1/B1 and A3/B3 and normal adjacent tissue A2/B2 and A4/B4). We conclude that histological and molecular data concur in more than three quarters of the cases and it appears that a substantial number of histologically normal tissues accumulate unphysiological levels of GALNTL5.

Accepted Article

In summary, GALNTL5 is a testicular protein that strongly accumulates in a substantial fraction of endometrium and thyroid cancer samples. We note that HPA reports weak RNA expression signals in brain samples and two out of three brain sections show GALNTL5 positive neuronal cells (www.proteinatlas.org; [47]). This pattern is reminiscent of the testis-brain specific gene class, although the reliability and biological relevance of non-testicular RNA/protein expression signals for GALNTL5 remain to be determined [6].

From our current study and earlier work published by others it emerges that RNA profiling alone is a sub-optimal approach to identifying testis-specific CT genes that are likely relevant for somatic cancer progression. Major caveats of transcriptomic approaches are that RNA signal intensities used to identify target genes are largely depend on the technologies and data normalization methods used and that transcribed CT gene mRNAs are not necessarily translated into physiologically relevant protein levels. Moreover, most current RNA profiling data from cancer tissues yields no information about how many and which cells in the sample express the target gene. Future work using improved single-cell RNA-Sequencing approaches will alleviate this critical issue. Large-scale protein profiling data obtained via IHC assays of testis, tumor and normal somatic samples might also be a promising method complementary to RNA-profiling, provided that the antibodies yield specific and reproducible results.

To further explore this protein-based approach, we selected two testis-specific genes for which we did not observe a typical CT gene pattern but that showed strong signals in cancer samples in HPA (www.proteinatlas.org; [47]). PDCL2 encodes a phosducin-like testis-specific protein [58]. Human PDCL2 is induced in the male germline and expressed at the RNA and protein level in male meiotic and post-meiotic germ cells (Figure 6D). Furthermore, its mRNA peaks in a frequent form of testicular cancer (seminoma, Figure 9A, B).

PDCL2 is a membrane protein that accumulates in round spermatids and in kidney cancer (www.proteinatlas.org; [47]). Using an HPA antibody (HPA048260) and a TMA for renal cancer (KD1503) we found that 92/100 (92%) of malignant samples show variable levels of staining (including 16 cases for which we observed strong signals) (Supplemental Figure S5C, in Figure 9C compare cancer samples C7-8/D7-8 and normal adjacent tissues C9/D9). While 31/50 (62%) of the NAT samples were positive, only two samples displayed strong staining (Supplemental Figure S5C, Figure 9C). This demonstrates that PDCL2 protein signals for renal cancer are reproducible using custom made and commercial TMAs [26]. The results also reveal that a

Accepted Article

substantial proportion of NAT samples appear to accumulate low levels of PDCL2. Since healthy kidney samples displayed no PDCL2 staining on our custom-made TMAs, while testicular germ cells are clearly marked, it is conceivable that some histologically normal tissues might already be precancerous at the molecular level (Figure 9D testis sample D6 and kidney samples A-C5). It is noteworthy that high expression of PDCL2 appears to be associated with decreased survival of kidney cancer patients (Supplemental Figure S6A).

C11orf42 mRNA is moderately expressed in adult testis, up-regulated in enriched round spermatids and although it is likely expressed in endometrium adenocarcinoma its RNA profile in control tissues does not mark it out as a *bona fide* CT gene (Figure 6E and Figure 10A, B). The protein is annotated as testis-specific and C11orf42 is detected in lung and thyroid cancer (www.proteinatlas.org; [47, 26]). Consistently, we observed cytoplasmic signals for C11orf42 in 3/40 (8%) thyroid cancer samples on a commercial TMA (TH481), while 0/8 of the normal controls showed cellular staining (Supplemental Figure S5D; Figure 10C and cancer sample A3 versus normal sample F3 in panel D).

Taken together, these results underline the robustness of combined RNA/protein-based methods to identify novel Cancer/Testis genes and highlight limits of approaches based on RNA profiling alone.

4. Discussion

We combined RNA/protein expression data from testis, male germ cells, normal controls and numerous somatic cancers to identify novel CT genes suitable for biomarker discovery and mechanistic analyses in the field of molecular oncogenesis.

4.1. The difficulty of identifying Cancer/Testis-specifically expressed genes

The male gonad is a complex organ and expresses the largest number of known genes among all tissues analysed so far, together with the brain [59, 18]. However, identifying bona fide testis-specific genes is a challenging task because somatic tissues that are used as negative controls in profiling studies are typically also composed of different cells types. When only a small subpopulation of cells in such a tissue expresses the testicular gene, its mRNAs may be diluted

Accepted Article

below the threshold level of detection, thereby yielding a false negative somatic control sample. This can lead to incoherent results with protein-based assays, especially immunohistochemistry, that detects signals in any cell population (or layer) of a given somatic organ. An analysis at the single cell-resolution level of human somatic and reproductive tissues will facilitate tackling this critical issue [60, 61, 15].

4.2. CT genes are promising candidates for prognostic biomarkers

CT genes represent a rich source for genes that confer oncogenic properties when abnormally expressed in somatic cancer cells [62], reviewed in [3]. A growing body of evidence links CT gene expression levels to unfavorable or favorable outcomes in the progression of a variety of somatic cancers, which underlines the clinical importance of CT genes as potential biomarkers and oncogenes [3]. Among core CT genes we identified 31 cases for which such data revealed various prognostic outcomes, including three MAGEA family members (for more details, see www.proteinatlas.org; [26]). Additional examples among the core CT genes are CMTM1 (signaling molecule) and SLC1A6 (amino acid transporter), which are unfavorable expression markers for pancreatic and urothelial cancer, while KHDRBS3 (RNA splicing) and GLUD2 (glutamate dehydrogenase) are favorable markers for kidney cancer. Interestingly, the type of prognosis appears to be dependent on the tissue that is affected, because LEMD1 (signaling molecule) is a favorable marker for ovarian cancer but an unfavorable one for pancreatic cancer (see Supplemental File S2 and Supplemental Figure S7 for the complete list of genes). Such dual-function genes have been found to be involved in cell cycle regulation; for review, see [63].

4.3. Novel CT genes may act as oncogenes or tumor suppressors

SPESP1 was associated with homologous recombination repair [64] (referenced in www.genomernai.org; [65]) and binds LYN, a tyrosine protein kinase important for cell proliferation and the response to DNA damage [66] (http://thebiogrid.org; [67] and www.nextprot.org; [68]. Mouse SPESP1 interacts with CENPC1, a centromere binding protein

Accepted Article

that plays a role in mitotic chromosome segregation [68, 69]; referenced in IntAct [70]. In summary, misexpression of SPESP1 may contribute to genetic instability and altered growth properties in somatic cancer cells.

GALNTL5 interacts with RHOU, a Rho-related GTP binding protein implicated in cancer cell migration (reviewed in [71]; [72, 73]) and TP53BP1, a protein involved in double-strand break repair, response to DNA damage and telomere dynamics [68, 74]; http://thebiogrid.org; [67]. This points to a potential role for GALNTL5 in cancer cell division and resistance to chemotherapy based on drugs that introduce DNA breaks, such as 5-fluorouracil and cisplatin [75].

A genome-wide RNAi screen identified C11orf42 as being important for normal mammary epithelial cell growth *in vitro* [76] (www.genomernai.org; [65]). In light of this potential role in cell division it is noteworthy that C11orf42 physically interacts with the protein transporter SNX5 [77] (IntAct [70]), which is expressed in the male germline (www.proteinatlas.org; [47]; www.germonline.org; [78]; https://rgv.genouest.org; [16]). Given that SNX5 is a negative prognostic marker for liver cancer and plays a role in promoting thyroid cancer progression by stabilizing growth factor receptors, C11orf42 may contribute to these pathological processes via its interaction with SNX5 [79, 26].

PDCL2 interacts with ACTRT1 and REST (RE1-silencing transcription factor) (http://thebiogrid.org; [67]). ACTRT1 is associated with sporadic basal cell carcinoma [80]. Mutations in REST predispose to Wilms tumor (the most common form of childhood renal cancer), suggesting that the gene acts as a tumor suppressor in this pediatric cancer [81]. We note that high REST expression correlates with increased survival in kidney cancer contrary to PDCL2, which shows the opposite effect (http://timer.cistrome.org; Supplemental Figure S5A,B). This raises the intriguing possibility that PDCL2 may act as a negative regulator of a renal tumor suppressor gene via direct protein-protein interaction with REST.

5. Conclusions

The accumulating evidence underlines that CT gene products, which have been touted as major targets for tumor neoantigen-based immunotherapies, are also interesting from an oncogenic perspective [82]. Further mechanistic studies of testicular proteins abnormally expressed in somatic cancer cells will help gain insight into molecular oncogenic processes. Such work may

Accepted Article

therefore facilitate efforts optimize existing treatments or even open up novel therapeutic opportunities.

6. Data accessibility

The entire dataset was generated with Affymetrix Human Genome U133 Plus 2.0 GeneChips (ThermoFisher). Expression data for human testis, testicular biopsies and enriched germ cells were described in reference [9]. Expression data for normal somatic control tissues were downloaded from the NCBI's Gene Omnibus (GEO; GSE7307, GSE6565, GSE11839) repository [10]. Cancer expression data produced by the Expression for Oncology (expO) project (www.intgen.org) were retrieved from GEO (GSE2109) and combined with two other datasets (GSE10802 and GSE6891).

7. Author contributions

SJ performed experiments and analysed data, FN analysed and interpreted data, CL interpreted data and provided reagents, RM and BJ contributed to study design, FC analysed data, MP conceived the study, interpreted data and wrote the manuscript. All authors contributed to the manuscript.

8. Acknowledgments

We thank Biosit's Histopathologie Haute Précision (H^2P^2) platform members. Some of our results are based upon data generated by the TCGA Research Network: www.cancer.gov/tcga. This work was funded by research grants of Fondation pour la Recherche Medicale (INE20071111109), Association pour la Recherche sur le Cancer (7985) and La Ligue Contre le Cancer committees CD29, CD35 and CD85 to M. Primig.

9. Conflict of interest

The authors declare no conflict of interest.

Accepted Article

10. References

1. Simpson AJ, Caballero OL, Jungbluth A, Chen YT & Old LJ (2005) Cancer/testis antigens, gametogenesis and cancer. Nat Rev Cancer 5, 615-625, doi: nrc1669 [pii] 10.1038/nrc1669.

2. Maxfield KE, Taus PJ, Corcoran K, Wooten J, Macion J, Zhou Y, Borromeo M, Kollipara RK, Yan J, Xie Y, Xie XJ & Whitehurst AW (2015) Comprehensive functional characterization of cancer-testis antigens defines obligate participation in multiple hallmarks of cancer. Nature communications 6, 8840, doi: 10.1038/ncomms9840.

3. Gibbs ZA & Whitehurst AW (2018) Emerging Contributions of Cancer/Testis Antigens to Neoplastic Behaviors. Trends Cancer 4, 701-712, doi: 10.1016/j.trecan.2018.08.005.

4. Gordeeva O (2018) Cancer-testis antigens: Unique cancer stem cell biomarkers and targets for cancer therapy. Semin Cancer Biol 53, 75-89, doi: 10.1016/j.semcancer.2018.08.006.

5. Almeida LG, Sakabe NJ, deOliveira AR, Silva MC, Mundstein AS, Cohen T, Chen YT, Chua R, Gurung S, Gnjatic S, Jungbluth AA, Caballero OL, Bairoch A, Kiesler E, White SL, Simpson AJ, Old LJ, Camargo AA & Vasconcelos AT (2009) CTdatabase: a knowledge-base of high-throughput and curated data on cancer-testis antigens. Nucleic Acids Res 37, D816-819, doi: 10.1093/nar/gkn673.

6. Hofmann O, Caballero OL, Stevenson BJ, Chen YT, Cohen T, Chua R, Maher CA, Panji S, Schaefer U, Kruger A, Lehvaslaiho M, Carninci P, Hayashizaki Y, Jongeneel CV, Simpson AJ, Old LJ & Hide W (2008) Genome-wide analysis of cancer/testis gene expression. Proc Natl Acad Sci U S A 105, 20422-20427, doi: 10.1073/pnas.0810777105.

7. Djureinovic D, Fagerberg L, Hallstrom B, Danielsson A, Lindskog C, Uhlen M & Ponten F (2014) The human testis-specific proteome defined by transcriptomics and antibody-based profiling. Mol Hum Reprod 20, 476-488, doi: 10.1093/molehr/gau018.

8. Djureinovic D, Hallstrom BM, Horie M, Mattsson JS, La Fleur L, Fagerberg L, Brunnstrom H, Lindskog C, Madjar K, Rahnenfuhrer J, Ekman S, Stahle E, Koyi H, Branden E, Edlund K, Hengstler JG, Lambe M, Saito A, Botling J, Ponten F, Uhlen M & Micke P (2016) Profiling cancer testis antigens in non-small-cell lung cancer. JCI Insight 1, e86837, doi: 10.1172/jci.insight.86837.

9. Chalmel F, Lardenois A, Evrard B, Mathieu R, Feig C, Demougin P, Gattiker A, Schulze W, Jegou B, Kirchhoff C & Primig M (2012) Global human tissue profiling and protein network

Accepted Article

analysis reveals distinct levels of transcriptional germline-specificity and identifies target genes for male infertility. Human reproduction 27, 3233-3248, doi: 10.1093/humrep/des301.

10. Clough E & Barrett T (2016) The Gene Expression Omnibus Database. Methods Mol Biol 1418, 93-110, doi: 10.1007/978-1-4939-3578-9_5.

11. Chalmel F, Rolland AD, Niederhauser-Wiederkehr C, Chung SS, Demougin P, Gattiker A, Moore J, Patard JJ, Wolgemuth DJ, Jegou B & Primig M (2007) The conserved transcriptome in human and rodent male gametogenesis. Proc Natl Acad Sci U S A 104, 8346-8351.

12. Chalmel F & Primig M (2008) The Annotation, Mapping, Expression and Network (AMEN) suite of tools for molecular systems biology. BMC Bioinformatics 9, 86.

13. Jegou B, Sankararaman S, Rolland AD, Reich D & Chalmel F (2017) Meiotic Genes Are Enriched in Regions of Reduced Archaic Ancestry. Mol Biol Evol 34, 1974-1980, doi: 10.1093/molbev/msx141.

14. Rolland AD, Evrard B, Darde TA, Le Beguec C, Le Bras Y, Bensalah K, Lavoue S, Jost B, Primig M, Dejucq-Rainsford N, Chalmel F & Jegou B (2019) RNA profiling of human testicular cells identifies syntenic lncRNAs associated with spermatogenesis. Human reproduction 34, 1278-1290, doi: 10.1093/humrep/dez063.

15. Wang M, Liu X, Chang G, Chen Y, An G, Yan L, Gao S, Xu Y, Cui Y, Dong J, Chen Y, Fan X, Hu Y, Song K, Zhu X, Gao Y, Yao Z, Bian S, Hou Y, Lu J, Wang R, Fan Y, Lian Y, Tang W, Wang Y, Liu J, Zhao L, Wang L, Liu Z, Yuan R, Shi Y, Hu B, Ren X, Tang F, Zhao XY & Qiao J (2018) Single-Cell RNA Sequencing Analysis Reveals Sequential Cell Fate Transition during Human Spermatogenesis. Cell Stem Cell 23, 599-614 e594, doi: 10.1016/j.stem.2018.08.007.

16. Darde TA, Lecluze E, Lardenois A, Stevant I, Alary N, Tuttelmann F, Collin O, Nef S, Jegou B, Rolland AD & Chalmel F (2019) The ReproGenomics Viewer: a multi-omics and cross-species resource compatible with single-cell studies for the reproductive science community. Bioinformatics 35, 3133-3139, doi: 10.1093/bioinformatics/btz047.

17. Darde TA, Sallou O, Becker E, Evrard B, Monjeaud C, Le Bras Y, Jegou B, Collin O, Rolland AD & Chalmel F (2015) The ReproGenomics Viewer: an integrative cross-species toolbox for the reproductive science community. Nucleic Acids Res 43, W109-116, doi: 10.1093/nar/gkv345.

Accepted Article

18. Uhlen M, Hallstrom BM, Lindskog C, Mardinoglu A, Ponten F & Nielsen J (2016) Transcriptomics resources of human tissues and organs. Molecular systems biology 12, 862, doi: 10.15252/msb.20155865.

19. Gao GF, Parker JS, Reynolds SM, Silva TC, Wang LB, Zhou W, Akbani R, Bailey M, Balu S, Berman BP, Brooks D, Chen H, Cherniack AD, Demchok JA, Ding L, Felau I, Gaheen S, Gerhard DS, Heiman DI, Hernandez KM, Hoadley KA, Jayasinghe R, Kemal A, Knijnenburg TA, Laird PW, Mensah MKA, Mungall AJ, Robertson AG, Shen H, Tarnuzzer R, Wang Z, Wyczalkowski M, Yang L, Zenklusen JC, Zhang Z, Genomic Data Analysis N, Liang H & Noble MS (2019) Before and After: Comparison of Legacy and Harmonized TCGA Genomic Data Commons' Data. Cell Syst 9, 24-34 e10, doi: 10.1016/j.cels.2019.06.006.

20. Petit FG, Kervarrec C, Jamin SP, Smagulova F, Hao C, Becker E, Jegou B, Chalmel F & Primig M (2015) Combining RNA and Protein Profiling Data with Network Interactions Identifies Genes associated with Spermatogenesis in Mouse and Human. Biology of reproduction, doi: 10.1095/biolreprod.114.126250.

21. Chalmel F, Lardenois A & Primig M (2007) Toward understanding the core meiotic transcriptome in mammals and its implications for somatic cancer. Ann N Y Acad Sci 1120, 1-15.

22. Sayers EW, Agarwala R, Bolton EE, Brister JR, Canese K, Clark K, Connor R, Fiorini N, Funk K, Hefferon T, Holmes JB, Kim S, Kimchi A, Kitts PA, Lathrop S, Lu Z, Madden TL, Marchler-Bauer A, Phan L, Schneider VA, Schoch CL, Pruitt KD & Ostell J (2019) Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 47, D23-D28, doi: 10.1093/nar/gky1069.

23. Bruggeman JW, Koster J, Lodder P, Repping S & Hamer G (2018) Massive expression of germ cell-specific genes is a hallmark of cancer and a potential target for novel treatment development. Oncogene 37, 5694-5700, doi: 10.1038/s41388-018-0357-2.

24. da Silva VL, Fonseca AF, Fonseca M, da Silva TE, Coelho AC, Kroll JE, de Souza JES, Stransky B, de Souza GA & de Souza SJ (2017) Genome-wide identification of cancer/testis genes and their association with prognosis in a pan-cancer analysis. Oncotarget 8, 92966-92977, doi: 10.18632/oncotarget.21715.

25. 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, Huo R, Wang Q, Xu L, Yang L, Huang X, Shen H, Sha J & Hu Z (2016) Systematic identification of genes with a cancer-testis expression pattern in 19 cancer types. Nature communications 7, 10499, doi: 10.1038/ncomms10499.

Accepted Article

26. Uhlen M, Zhang C, Lee S, Sjostedt E, Fagerberg L, Bidkhori G, Benfeitas R, Arif M, Liu Z, Edfors F, Sanli K, von Feilitzen K, Oksvold P, Lundberg E, Hober S, Nilsson P, Mattsson J, Schwenk JM, Brunnstrom H, Glimelius B, Sjoblom T, Edqvist PH, Djureinovic D, Micke P, Lindskog C, Mardinoglu A & Ponten F (2017) A pathology atlas of the human cancer transcriptome. Science 357, doi: 10.1126/science.aan2507.

27. Delic S, Thuy A, Schulze M, Proescholdt MA, Dietrich P, Bosserhoff AK & Riemenschneider MJ (2015) Systematic investigation of CMTM family genes suggests relevance to glioblastoma pathogenesis and CMTM1 and CMTM3 as priority targets. Genes Chromosomes Cancer 54, 433-443, doi: 10.1002/gcc.22255.

28. Paret C, Simon P, Vormbrock K, Bender C, Kolsch A, Breitkreuz A, Yildiz O, Omokoko T, Hubich-Rau S, Hartmann C, Hacker S, Wagner M, Roldan DB, Selmi A, Tureci O & Sahin U (2015) CXorf61 is a target for T cell based immunotherapy of triple-negative breast cancer. Oncotarget 6, 25356-25367, doi: 10.18632/oncotarget.4516.

29. Jain SU, Do TJ, Lund PJ, Rashoff AQ, Diehl KL, Cieslik M, Bajic A, Juretic N, Deshmukh S, Venneti S, Muir TW, Garcia BA, Jabado N & Lewis PW (2019) PFA ependymoma-associated protein EZHIP inhibits PRC2 activity through a H3 K27M-like mechanism. Nature communications 10, 2146, doi: 10.1038/s41467-019-09981-6.

30. Wang H, Chen Y, Han J, Meng Q, Xi Q, Wu G & Zhang B (2016) DCAF4L2 promotes colorectal cancer invasion and metastasis via mediating degradation of NFkappab negative regulator PPM1B. Am J Transl Res 8, 405-418.

31. Matsumoto Y, Itou J, Sato F & Toi M (2018) SALL4 - KHDRBS3 network enhances stemness by modulating CD44 splicing in basal-like breast cancer. Cancer Med 7, 454-462, doi: 10.1002/cam4.1296.

32. Sasahira T, Kurihara M, Nakashima C, Kirita T & Kuniyasu H (2016) LEM domain containing 1 promotes oral squamous cell carcinoma invasion and endothelial transmigration. Br J Cancer 115, 52-58, doi: 10.1038/bjc.2016.167.

33. Xie K, Zhang K, Kong J, Wang C, Gu Y, Liang C, Jiang T, Qin N, Liu J, Guo X, Huo R, Liu M, Ma H, Dai J & Hu Z (2018) Cancer-testis gene PIWIL1 promotes cell proliferation, migration, and invasion in lung adenocarcinoma. Cancer Med 7, 157-166, doi: 10.1002/cam4.1248.

34. Huo S, Du W, Shi P, Si Y & Zhao S (2015) The role of spermatogenesis-associated protein 6 in testicular germ cell tumors. Int J Clin Exp Pathol 8, 9119-9125.

Accepted Article

35. Yao J, Caballero OL, Yung WK, Weinstein JN, Riggins GJ, Strausberg RL & Zhao Q (2014) Tumor subtype-specific cancer-testis antigens as potential biomarkers and immunotherapeutic targets for cancers. Cancer Immunol Res 2, 371-379, doi: 10.1158/2326-6066.CIR-13-0088.

36. Scanlan MJ, Welt S, Gordon CM, Chen YT, Gure AO, Stockert E, Jungbluth AA, Ritter G, Jager D, Jager E, Knuth A & Old LJ (2002) Cancer-related serological recognition of human colon cancer: identification of potential diagnostic and immunotherapeutic targets. Cancer Res 62, 4041-4047.

37. Tang L, Chen F, Pang EJ, Zhang ZQ, Jin BW & Dong WF (2015) MicroRNA-182 inhibits proliferation through targeting oncogenic ANUBL1 in gastric cancer. Oncol Rep 33, 1707-1716, doi: 10.3892/or.2015.3798.

38. Ji S, Zhang W, Zhang X, Hao C, Hao A, Gao Q, Zhang H, Sun J & Hao J (2016) Sohlh2 suppresses epithelial to mesenchymal transition in breast cancer via downregulation of IL-8. Oncotarget 7, 49411-49424, doi: 10.18632/oncotarget.10355.

39. Fok KL, Chung CM, Yi SQ, Jiang X, Sun X, Chen H, Chen YC, Kung HF, Tao Q, Diao R, Chan H, Zhang XH, Chung YW, Cai Z & Chang Chan H (2012) STK31 maintains the undifferentiated state of colon cancer cells. Carcinogenesis 33, 2044-2053, doi: 10.1093/carcin/bgs246.

40. Shuptrine CW, Ajina R, Fertig EJ, Jablonski SA, Kim Lyerly H, Hartman ZC & Weiner LM (2017) An unbiased in vivo functional genomics screening approach in mice identifies novel tumor cell-based regulators of immune rejection. Cancer Immunol Immunother 66, 1529-1544, doi: 10.1007/s00262-017-2047-2.

41. Holland AJ & Cleveland DW (2012) The deubiquitinase USP44 is a tumor suppressor that protects against chromosome missegregation. The Journal of clinical investigation 122, 4325-4328, doi: 10.1172/JCI66420.

42. Ma D, Yang J, Wang Y, Huang X, Du G & Zhou L (2017) Whole exome sequencing identified genetic variations in Chinese hemangioblastoma patients. Am J Med Genet A 173, 2605-2613, doi: 10.1002/ajmg.a.38350.

43. Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q, Li B & Liu XS (2020) TIMER2.0 for analysis of tumor-infiltrating immune cells. Nucleic Acids Res 48, W509-W514, doi: 10.1093/nar/gkaa407.

Accepted Article

44. Zhang T & Zarkower D (2017) DMRT proteins and coordination of mammalian spermatogenesis. Stem Cell Res 24, 195-202, doi: 10.1016/j.scr.2017.07.026.

45. Kanetsky PA, Mitra N, Vardhanabhuti S, Vaughn DJ, Li M, Ciosek SL, Letrero R, D'Andrea K, Vaddi M, Doody DR, Weaver J, Chen C, Starr JR, Hakonarson H, Rader DJ, Godwin AK, Reilly MP, Schwartz SM & Nathanson KL (2011) A second independent locus within DMRT1 is associated with testicular germ cell tumor susceptibility. Human molecular genetics 20, 3109-3117, doi: 10.1093/hmg/ddr207.

46. Turnbull C, Rapley EA, Seal S, Pernet D, Renwick A, Hughes D, Ricketts M, Linger R, Nsengimana J, Deloukas P, Huddart RA, Bishop DT, Easton DF, Stratton MR, Rahman N & Collaboration UKTC (2010) Variants near DMRT1, TERT and ATF7IP are associated with testicular germ cell cancer. Nat Genet 42, 604-607, doi: 10.1038/ng.607.

47. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson A, Kampf C, Sjostedt E, Asplund A, Olsson I, Edlund K, Lundberg E, Navani S, Szigyarto CA, Odeberg J, Djureinovic D, Takanen JO, Hober S, Alm T, Edqvist PH, Berling H, Tegel H, Mulder J, Rockberg J, Nilsson P, Schwenk JM, Hamsten M, von Feilitzen K, Forsberg M, Persson L, Johansson F, Zwahlen M, von Heijne G, Nielsen J & Ponten F (2015) Proteomics. Tissue-based map of the human proteome. Science 347, 1260419, doi: 10.1126/science.1260419.

48. Miao Y, Cui L, Chen Z & Zhang L (2016) Gene expression profiling of DMU-212-induced apoptosis and anti-angiogenesis in vascular endothelial cells. Pharm Biol 54, 660-666, doi: 10.3109/13880209.2015.1071414.

49. Bullon P & Navarro JM (2017) Inflammasome as a Key Pathogenic Mechanism in Endometriosis. Curr Drug Targets 18, 997-1002, doi: 10.2174/1389450117666160709013850.

50. Yang F, De La Fuente R, Leu NA, Baumann C, McLaughlin KJ & Wang PJ (2006) Mouse SYCP2 is required for synaptonemal complex assembly and chromosomal synapsis during male meiosis. The Journal of cell biology 173, 497-507, doi: 10.1083/jcb.200603063.

51. Schilit SLP, Menon S, Friedrich C, Kammin T, Wilch E, Hanscom C, Jiang S, Kliesch S, Talkowski ME, Tuttelmann F, MacQueen AJ & Morton CC (2020) SYCP2 Translocation-Mediated Dysregulation and Frameshift Variants Cause Human Male Infertility. Am J Hum Genet 106, 41-57, doi: 10.1016/j.ajhg.2019.11.013.

52. Wu C & Tuo Y (2019) SYCP2 expression is a novel prognostic biomarker in luminal A/B breast cancer. Future Oncol 15, 817-826, doi: 10.2217/fon-2018-0821.

Accepted Article

53. Fujihara Y, Murakami M, Inoue N, Satouh Y, Kaseda K, Ikawa M & Okabe M (2010) Sperm equatorial segment protein 1, SPESP1, is required for fully fertile sperm in mouse. Journal of cell science 123, 1531-1536, doi: 10.1242/jcs.067363.

54. Ito C, Yamatoya K, Yoshida K, Fujimura L, Sugiyama H, Suganami A, Tamura Y, Hatano M, Miyado K & Toshimori K (2018) Deletion of Eqtn in mice reduces male fertility and sperm-egg adhesion. Reproduction, doi: 10.1530/REP-18-0394.

55. Wolkowicz MJ, Digilio L, Klotz K, Shetty J, Flickinger CJ & Herr JC (2008) Equatorial segment protein (ESP) is a human alloantigen involved in sperm-egg binding and fusion. Journal of andrology 29, 272-282, doi: 10.2164/jandrol.106.000604.

56. Takasaki N, Tachibana K, Ogasawara S, Matsuzaki H, Hagiuda J, Ishikawa H, Mochida K, Inoue K, Ogonuki N, Ogura A, Noce T, Ito C, Toshimori K & Narimatsu H (2014) A heterozygous mutation of GALNTL5 affects male infertility with impairment of sperm motility. Proc Natl Acad Sci U S A 111, 1120-1125, doi: 10.1073/pnas.1310777111.

57. Hagiuda J, Takasaki N, Oya M, Ishikawa H & Narimatsu H (2019) Mutation of GALNTL5 gene identified in patients diagnosed with asthenozoospermia. Hum Fertil (Camb), 1-8, doi: 10.1080/14647273.2018.1562239.

58. Lou X, Bao R, Zhou CZ & Chen Y (2009) Structure of the thioredoxin-fold domain of human phosducin-like protein 2. Acta Crystallogr Sect F Struct Biol Cryst Commun 65, 67-70, doi: 10.1107/S1744309108037342.

59. Consortium GT (2015) Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science 348, 648-660, doi: 10.1126/science.1262110.

60. Bagnoli JW, Wange LE, Janjic A & Enard W (2019) Studying Cancer Heterogeneity by Single-Cell RNA Sequencing. Methods Mol Biol 1956, 305-319, doi: 10.1007/978-1-4939-9151-8_14.

61. Guo J, Grow EJ, Mlcochova H, Maher GJ, Lindskog C, Nie X, Guo Y, Takei Y, Yun J, Cai L, Kim R, Carrell DT, Goriely A, Hotaling JM & Cairns BR (2018) The adult human testis transcriptional cell atlas. Cell research 28, 1141-1157, doi: 10.1038/s41422-018-0099-2.

62. Maheswaran E, Pedersen CB, Ditzel HJ & Gjerstorff MF (2015) Lack of ADAM2, CALR3 and SAGE1 Cancer/Testis Antigen Expression in Lung and Breast Cancer. PloS one 10, e0134967, doi: 10.1371/journal.pone.0134967.

Accepted Article

63. Lou X, Zhang J, Liu S, Xu N & Liao DJ (2014) The other side of the coin: the tumor-suppressive aspect of oncogenes and the oncogenic aspect of tumor-suppressive genes, such as those along the CCND-CDK4/6-RB axis. 13, 1677-1693, doi: 10.4161/cc.29082.

64. Slabicki M, Theis M, Krastev DB, Samsonov S, Mundwiller E, Junqueira M, Paszkowski-Rogacz M, Teyra J, Heninger AK, Poser I, Prieur F, Truchetto J, Confavreux C, Marelli C, Durr A, Camdessanche JP, Brice A, Shevchenko A, Pisabarro MT, Stevanin G & Buchholz F (2010) A genome-scale DNA repair RNAi screen identifies SPG48 as a novel gene associated with hereditary spastic paraplegia. PLoS Biol 8, e1000408, doi: 10.1371/journal.pbio.1000408.

65. Schmidt EE, Pelz O, Buhlmann S, Kerr G, Horn T & Boutros M (2013) GenomeRNAi: a database for cell-based and in vivo RNAi phenotypes, 2013 update. Nucleic Acids Res 41, D1021-1026, doi: 10.1093/nar/gks1170.

66. Zhu J, Larman HB, Gao G, Somwar R, Zhang Z, Laserson U, Ciccia A, Pavlova N, Church G, Zhang W, Kesari S & Elledge SJ (2013) Protein interaction discovery using parallel analysis of translated ORFs (PLATO). Nature biotechnology 31, 331-334, doi: 10.1038/nbt.2539.

67. Oughtred R, Stark C, Breitkreutz BJ, Rust J, Boucher L, Chang C, Kolas N, O'Donnell L, Leung G, McAdam R, Zhang F, Dolma S, Willems A, Coulombe-Huntington J, Chatr-Aryamontri A, Dolinski K & Tyers M (2019) The BioGRID interaction database: 2019 update. Nucleic Acids Res 47, D529-D541, doi: 10.1093/nar/gky1079.

68. Duek P, Gateau A, Bairoch A & Lane L (2018) Exploring the Uncharacterized Human Proteome Using neXtProt. Journal of proteome research 17, 4211-4226, doi: 10.1021/acs.jproteome.8b00537.

69. Kim J, Ishiguro K, Nambu A, Akiyoshi B, Yokobayashi S, Kagami A, Ishiguro T, Pendas AM, Takeda N, Sakakibara Y, Kitajima TS, Tanno Y, Sakuno T & Watanabe Y (2015) Meikin is a conserved regulator of meiosis-I-specific kinetochore function. Nature 517, 466-471, doi: 10.1038/nature14097.

70. Orchard S, Ammari M, Aranda B, Breuza L, Briganti L, Broackes-Carter F, Campbell NH, Chavali G, Chen C, del-Toro N, Duesbury M, Dumousseau M, Galeota E, Hinz U, Iannuccelli M, Jagannathan S, Jimenez R, Khadake J, Lagreid A, Licata L, Lovering RC, Meldal B, Melidoni AN, Milagros M, Peluso D, Perfetto L, Porras P, Raghunath A, Ricard-Blum S, Roechert B, Stutz A, Tognolli M, van Roey K, Cesareni G & Hermjakob H (2014) The MIntAct project--IntAct as a common curation platform for 11 molecular interaction databases. Nucleic Acids Res 42, D358-363, doi: 10.1093/nar/gkt1115.

Accepted Article

71. Faure S & Fort P (2015) Atypical RhoV and RhoU GTPases control development of the neural crest. Small GTPases 6, 174-177, doi: 10.1080/21541248.2015.1025943.

72. Canovas Nunes S, Manzoni M, Pizzi M, Mandato E, Carrino M, Quotti Tubi L, Zambello R, Adami F, Visentin A, Barila G, Trentin L, Manni S, Neri A, Semenzato G & Piazza F (2018) The small GTPase RhoU lays downstream of JAK/STAT signaling and mediates cell migration in multiple myeloma. Blood Cancer J 8, 20, doi: 10.1038/s41408-018-0053-z.

73. De Piano M, Manuelli V, Zadra G, Otte J, Edqvist PD, Ponten F, Nowinski S, Niaouris A, Grigoriadis A, Loda M, Van Hemelrijck M & Wells CM (2020) Lipogenic signalling modulates prostate cancer cell adhesion and migration via modification of Rho GTPases. Oncogene, doi: 10.1038/s41388-020-1243-2.

74. Woods NT, Mesquita RD, Sweet M, Carvalho MA, Li X, Liu Y, Nguyen H, Thomas CE, Iversen ES, Jr., Marsillac S, Karchin R, Koomen J & Monteiro AN (2012) Charting the landscape of tandem BRCT domain-mediated protein interactions. Sci Signal 5, rs6, doi: 10.1126/scisignal.2002255.

75. Swift LH & Golsteyn RM (2014) Genotoxic anti-cancer agents and their relationship to DNA damage, mitosis, and checkpoint adaptation in proliferating cancer cells. Int J Mol Sci 15, 3403-3431, doi: 10.3390/ijms15033403.

76. Burleigh A, McKinney S, Brimhall J, Yap D, Eirew P, Poon S, Ng V, Wan A, Prentice L, Annab L, Barrett JC, Caldas C, Eaves C & Aparicio S (2015) A co-culture genome-wide RNAi screen with mammary epithelial cells reveals transmembrane signals required for growth and differentiation. Breast Cancer Res 17, 4, doi: 10.1186/s13058-014-0510-y.

77. Huttlin EL, Bruckner RJ, Paulo JA, Cannon JR, Ting L, Baltier K, Colby G, Gebreab F, Gygi MP, Parzen H, Szpyt J, Tam S, Zarraga G, Pontano-Vaites L, Swarup S, White AE, Schweppe DK, Rad R, Erickson BK, Obar RA, Guruharsha KG, Li K, Artavanis-Tsakonas S, Gygi SP & Harper JW (2017) Architecture of the human interactome defines protein communities and disease networks. Nature 545, 505-509, doi: 10.1038/nature22366.

78. Lardenois A, Gattiker A, Collin O, Chalmel F & Primig M (2010) GermOnline 4.0 is a genomics gateway for germline development, meiosis and the mitotic cell cycle. Database (Oxford) 2010, baq030, doi: baq030 [pii]
10.1093/database/baq030.

79. Jitsukawa S, Kamekura R, Kawata K, Ito F, Sato A, Matsumiya H, Nagaya T, Yamashita K, Kubo T, Kikuchi T, Sato N, Hasegawa T, Kiyonari H, Mukumoto Y, Takano KI, Himi T &

Accepted Article

Ichimiya S (2017) Loss of sorting nexin 5 stabilizes internalized growth factor receptors to promote thyroid cancer progression. J Pathol 243, 342-353, doi: 10.1002/path.4951.

80. Bal E, Park HS, Belaid-Choucair Z, Kayserili H, Naville M, Madrange M, Chiticariu E, Hadj-Rabia S, Cagnard N, Kuonen F, Bachmann D, Huber M, Le Gall C, Cote F, Hanein S, Rosti RO, Aslanger AD, Waisfisz Q, Bodemer C, Hermine O, Morice-Picard F, Labeille B, Caux F, Mazereeuw-Hautier J, Philip N, Levy N, Taieb A, Avril MF, Headon DJ, Gyapay G, Magnaldo T, Fraitag S, Crollius HR, Vabres P, Hohl D, Munnich A & Smahi A (2017) Mutations in ACTRT1 and its enhancer RNA elements lead to aberrant activation of Hedgehog signaling in inherited and sporadic basal cell carcinomas. Nature medicine 23, 1226-1233, doi: 10.1038/nm.4368.

81. Mahamdallie SS, Hanks S, Karlin KL, Zachariou A, Perdeaux ER, Ruark E, Shaw CA, Renwick A, Ramsay E, Yost S, Elliott A, Birch J, Capra M, Gray J, Hale J, Kingston J, Levitt G, McLean T, Sheridan E, Renwick A, Seal S, Stiller C, Sebire N, Westbrook TF & Rahman N (2015) Mutations in the transcriptional repressor REST predispose to Wilms tumor. Nat Genet 47, 1471-1474, doi: 10.1038/ng.3440.

82. Gjerstorff MF, Andersen MH & Ditzel HJ (2015) Oncogenic cancer/testis antigens: prime candidates for immunotherapy. Oncotarget 6, 15772-15787, doi: 10.18632/oncotarget.4694.

11. Supporting information

Supplemental File S1. Searchable annotation and expression data.

Supplemental File S2. Core CT gene annotation, cancer prognosis and RNA-Sequencing data.

Supplemental Figure S1. Detection pattern of testicular proteins in somatic tissues.

Supplemental Figure S2. Core CT gene expression in testis and male germ cells.

Supplemental Figure S3. TCGA expression data.

Supplemental Figure S4. Kaplan-Meyer (KM) plots for CT genes.

Supplemental Figure S5. Commercial- and custom cancer TMA sample annotation.

Supplemental Figure S6. Kaplan-Meyer (KM) plot for PDCL2 and REST.

Supplemental Figure S7. Gene expression/cancer prognosis matrix.

Accepted Article

12. Figure legends

Figure 1. RNA expression data analysis procedure. (A) The experimental protocol and analysis procedure for CT gene screening are shown as a flow chart. Data from 2998 human U133 Plus 2.0 GeneChips were processed and normalized as shown. Probeset numbers are indicated for filtration steps and corresponding gene numbers are shown in parentheses. (B) A box plot of log2 intensity values indicates the median signal (shown as a red line) and values that correspond to expression confidence levels (not expressed, low expression, expressed and highly expressed) as well as the percentiles.

Figure 2. CT gene expression in testis, tumor and normal somatic samples. False-color heatmaps are shown that display the log2-transformed expression levels for testicular genes in normal somatic controls (left panel) and somatic cancers (right panel). Each line corresponds to a probeset and each column to a sample. The samples are clustered according to their expression pattern as shown by dendrograms at the top and to the left. Sample annotation categories are shown at the bottom. A color coded scale bar (red, and blue indicate high and low expression, respectively) for log2 values is given to the right.

Figure 3. CT gene expression profiles. A pie chart shows the distribution of CT genes referenced in CTDatabase (http://www.cta.lncc.br) among the expression classes defined by our analysis as indicated at the bottom. We employed probe sets for this analysis to avoid multiple allocations of genes to different classes. Percentages are rounded to the nearest integer.

Figure 4. Core CT gene expression in testis, somatic tissues and cancer. (A) A color-coded heatmap is shown for RNA-Sequencing data obtained with samples as indicated at the bottom. Each line corresponds to a gene, each column corresponds to a sample. Gene symbols are shown to the right. Genes and samples are grouped together using the Euclidian clustering algorithm; Dendrograms are shown at the top and to the left. Red and blue represent high and low expression levels. A scale is given at the bottom. (B) A color coded heatmap is shown for RNA-Sequencing

Accepted Article

data obtained with cancer samples as indicated at the bottom. Each line corresponds to a gene, each column corresponds to a sample. Gene symbols are shown to the right. Genes and samples are grouped together using the Euclidian clustering algorithm; Dendrograms are shown at the top and to the left. Red and blue represent high and low expression levels. A scale is given at the bottom.

Figure 5. CT gene expression profiles in testis, controls and cancer. (A-E) Bar diagrams display the log2-transformed GeneChip expression signals (y-axis) obtained with total testis samples and enriched germ cells (Testis & GC), controls (normal tissues) and malign tumor samples (cancer tissues; X-axis). A red line indicates the expression median. The target gene names are indicated at the top right corners. A box plot is shown like in Figure 1B. Detailed sample annotation is available in the searchable Supplemental File S1.

Figure 6. Testicular single cell CT gene expression profiles. (A) A schematic scatter plot of color-coded and annotated testicular cells is shown. A legend is given to the right. Replicates are numbered. (B-E) Scatter plots of single-cell expression data from testicular samples are shown for four CT genes. Legends represent expression levels in cell types as indicated (shades of red and grey indicate high and low expression, respectively).

Figure 7. SPESP1 RNA/protein expression in skin cancer versus control samples. (A) A bar diagram is shown for the expression data like in Figure 5. (B) Immunohistochemical staining is shown for SPESP1 on human adult testicular sections; a red rectangle marks an enlarged region given to the right. Scale bars 50 µm and 20µm. (C) An image of the hybridized TMA is shown. Samples are identified by numbers for columns and letters for rows. The TMA identifier and the cancer type are indicated at the bottom. Scale bar 5mm. (D) Enlarged images of typical skin cancer samples showing strong (A7-A8) or no (C7-C8) staining are shown at the top. Two normal skin controls (J7-J8) are shown at the bottom. Scale bar 500µm. A red rectangle marks an enlarged section given to the right. Scale bar 100µm.

Figure 8. GALNTL5 RNA/protein expression in endometrial and thyroid cancer versus control samples. (A-B) A bar diagram is shown for RNA expression data and an image of an

Accepted Article

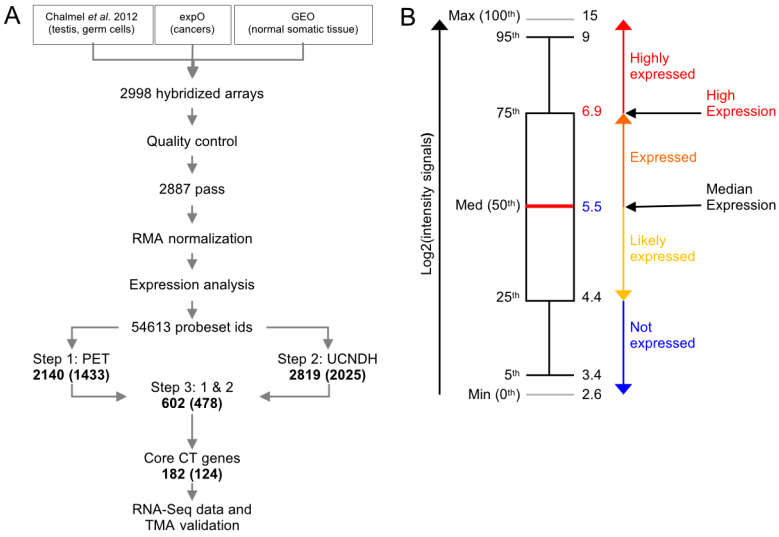
immunohistochemical assay is given. Scale bars 50 μm and 20μm. (C-E). Images and schematics of hybridized cancer TMAs are given like in Figure 7. Scale bars 5mm for the global view and 1mm for the detailed view.

Figure 9. PDCL2 RNA/protein expression in kidney cancer versus control samples. (A-B) A bar diagram is shown for RNA expression data and an image of an immunohistochemical assay is given. Scale bars 50 μm and 20μm. (C) An image of the hybridized TMA is shown. Scale bar 5mm. Tumor and NAT samples are identified by numbers for columns and letters for rows. The TMA identifier and the cancer type are indicated at the bottom. Scale bar 1mm. (D) An image of a custom-made TMA is given. Scale bar 5mm. The small rectangle over sample D6 identifies one of three testicular samples for which an enlarged images is shown at the bottom left side. Scale bar 100μm. The large rectangle identifies three normal kidney samples (A5-C5); an enlarged image is shown at the bottom right side. Scale bar 1mm.

Figure 10. C11orf42 RNA/protein expression in thyroid cancer versus control samples. (A-B) A bar diagram is shown for RNA expression data and an image of an immunohistochemical assay is given. Scale bars 50 μm and 20μm. (C) An image of the hybridized TMA is shown. Samples are identified by numbers for columns and letters for rows. The TMA identifier and the cancer type are indicated at the bottom. Scale bar 5mm. (D) Enlarged images of a positive cancer sample (A3) and a negative control sample (F3) are given at the bottom. Scale bar 500μm.

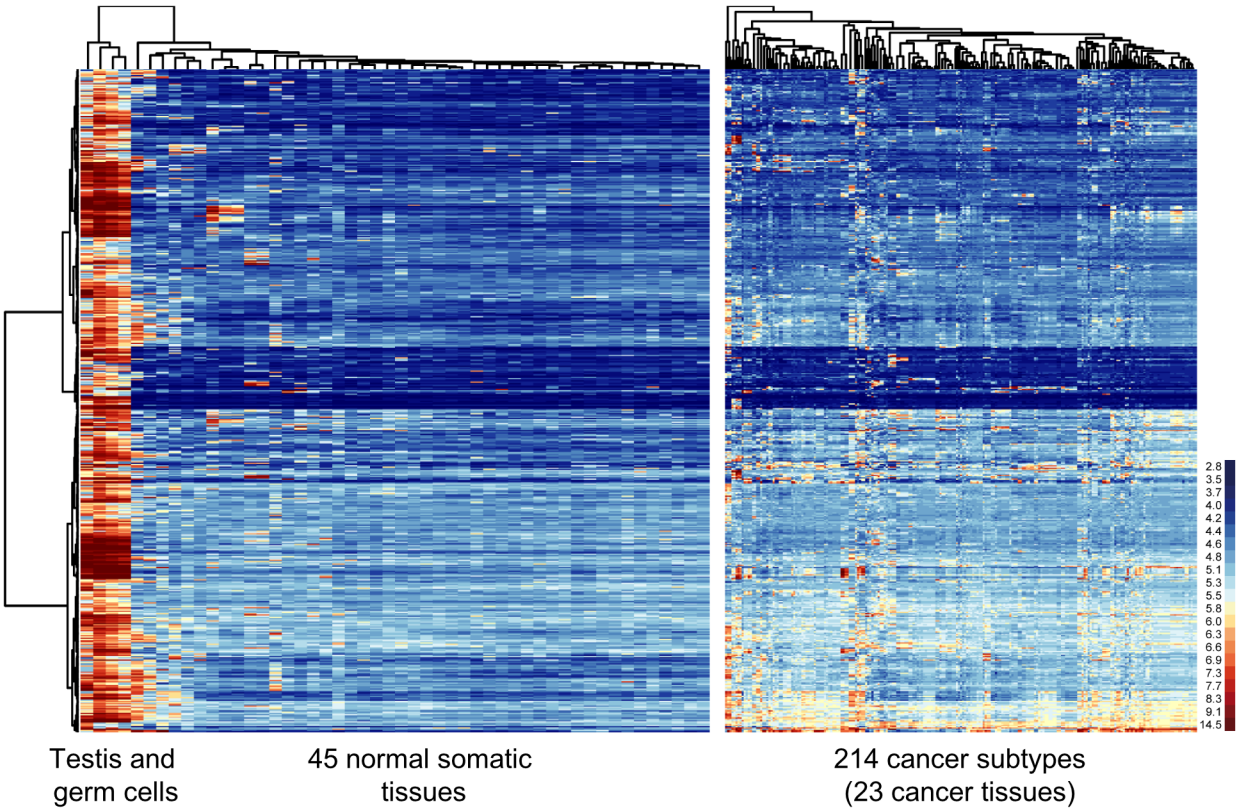
Accepted Article

Accepted Article



mol2_12900_f1.tiff

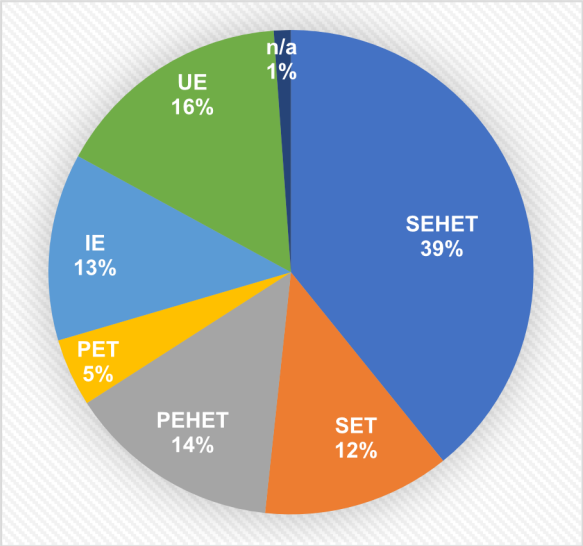
Accepted Article



mol2_12900_f2.tiff

Accepted Article

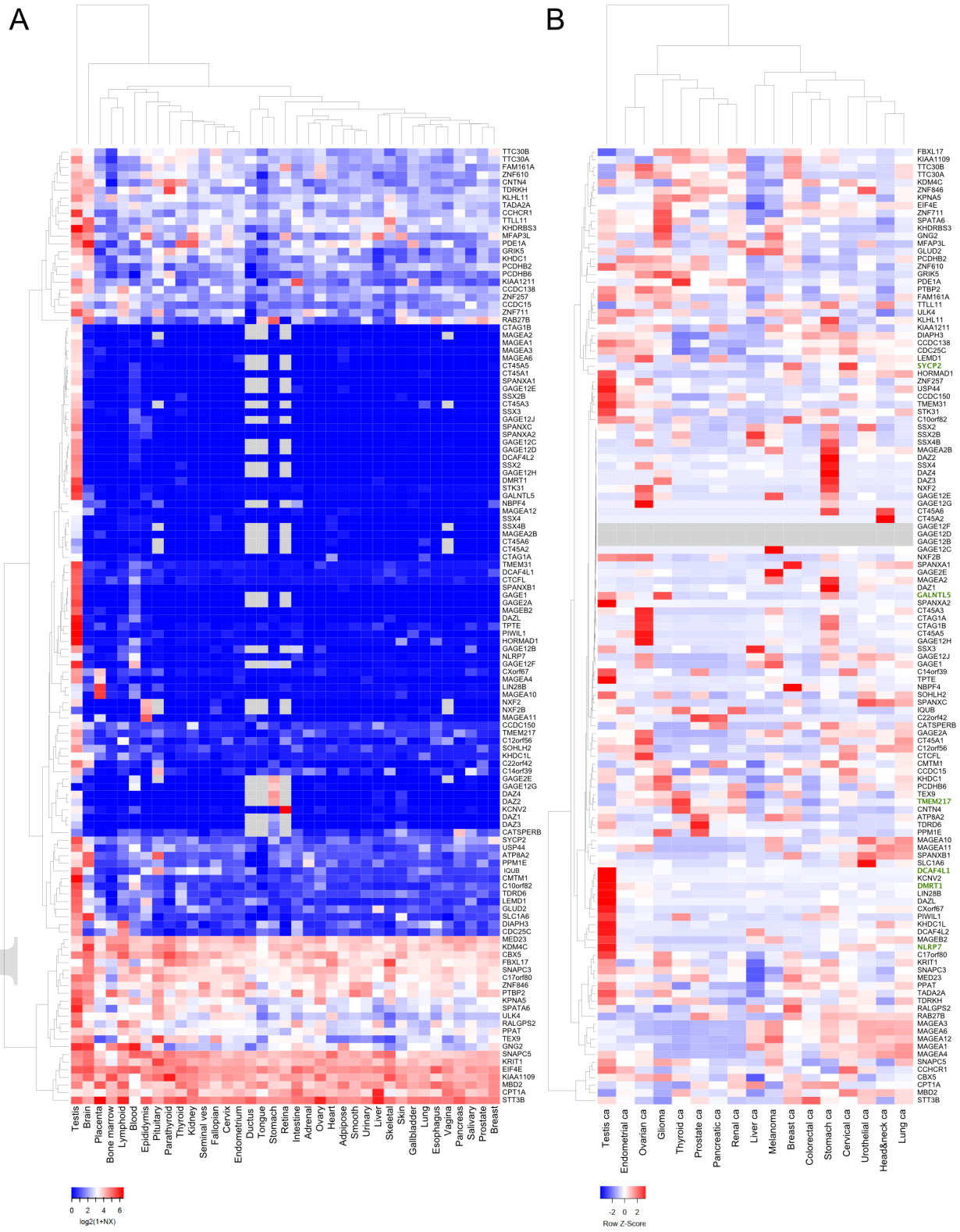
Referenced CTgenes



SET: specifically expressed in testis
SEHET: specifically and highly expressed in testis
PET: preferentially expressed in testis
PEHET: preferentially and highly expressed in testis
IE: intermediate expression
UE: ubiquitous expression
n/a: not available

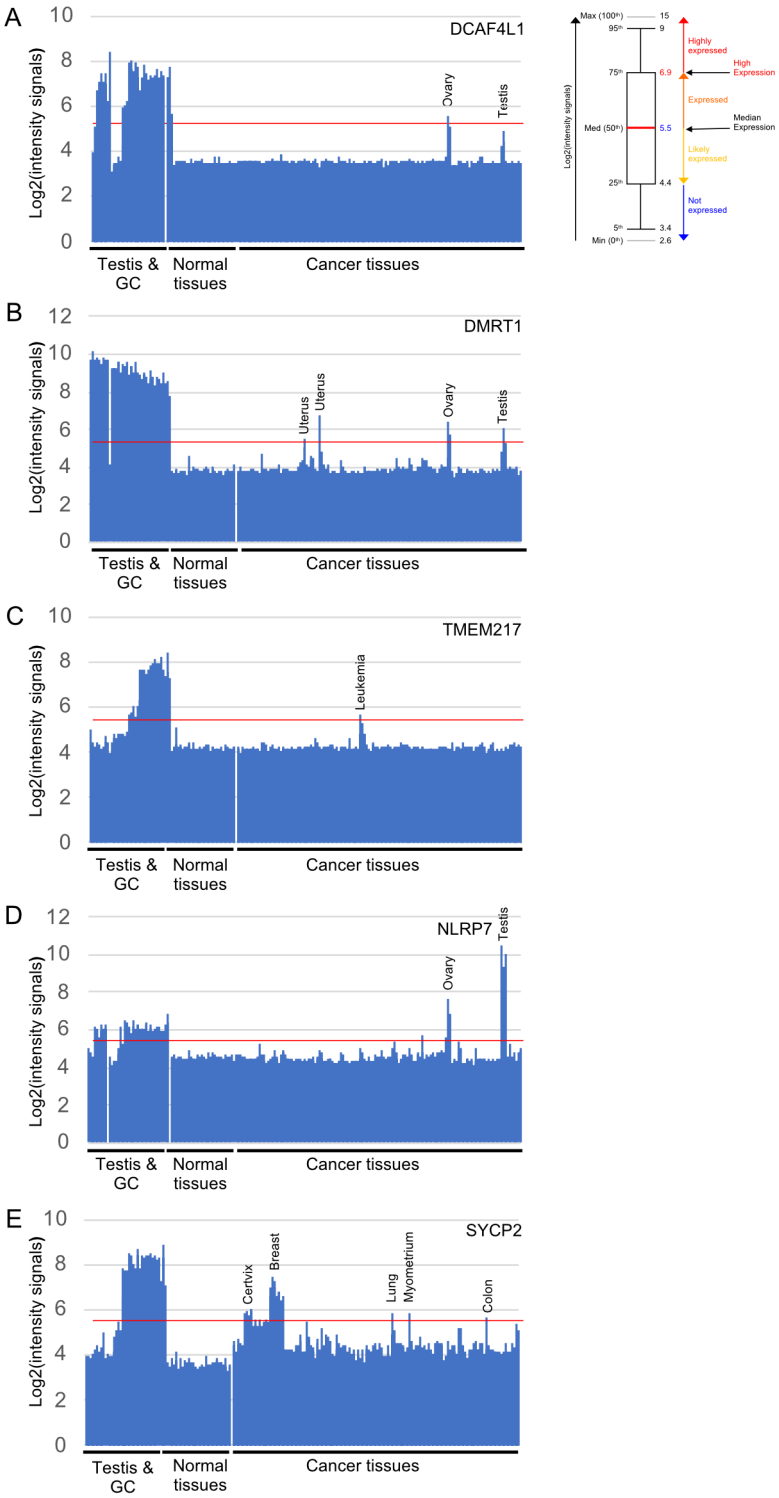
mol2_12900_f3.tiff

Accepted Article



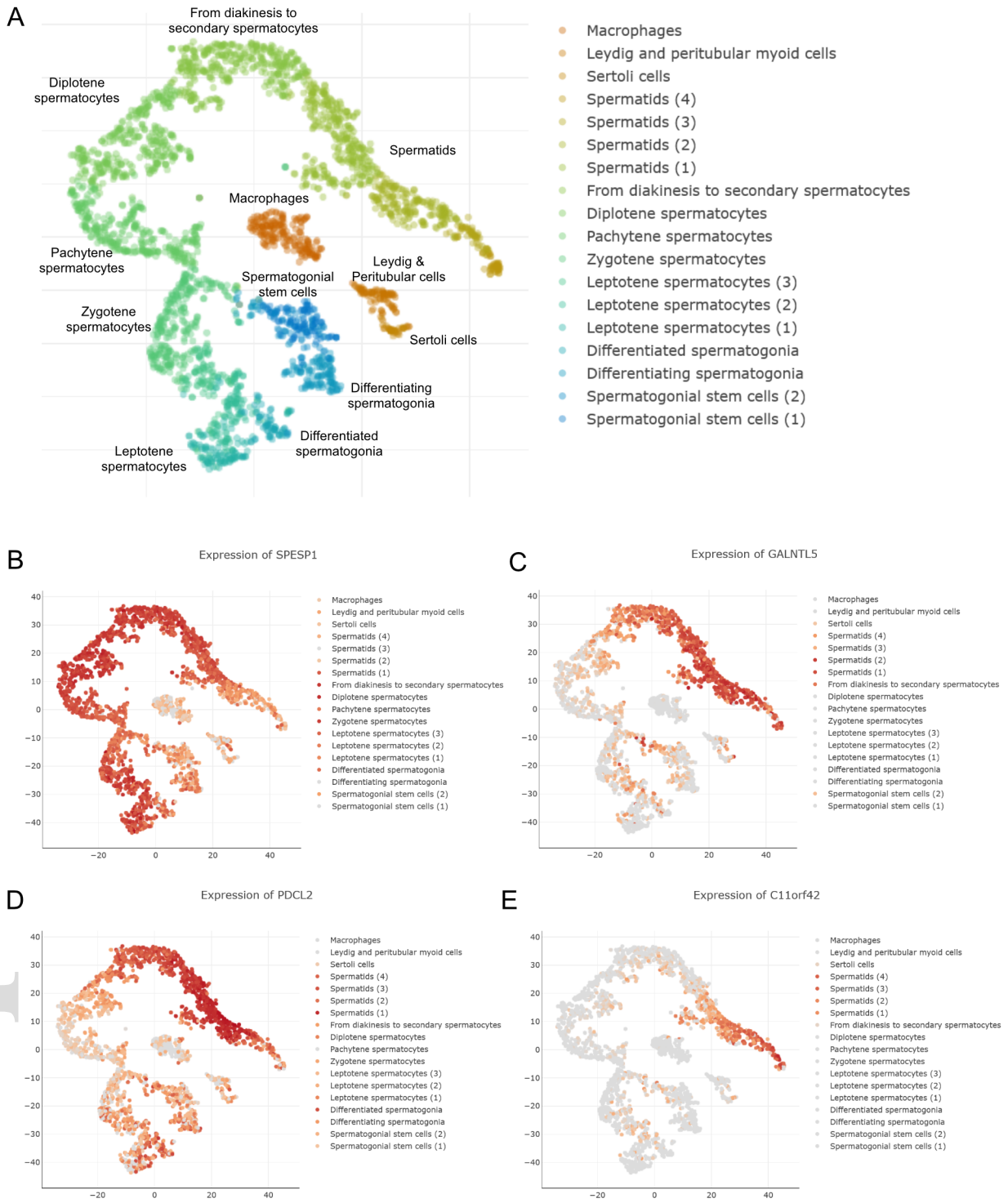
mol2_12900_f4.tiff

Accepted Article



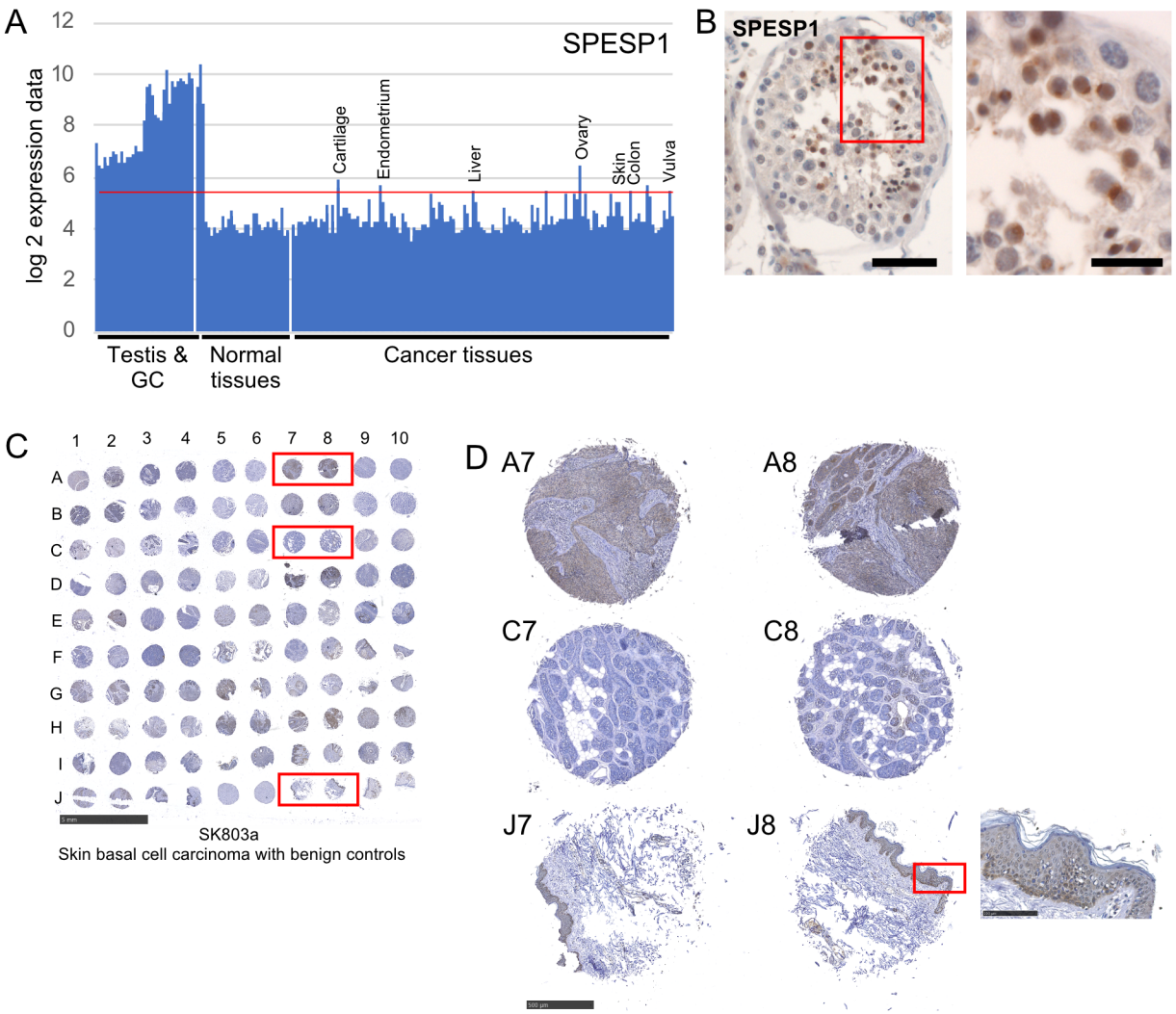
mol2_12900_f5.tiff

Accepted Article



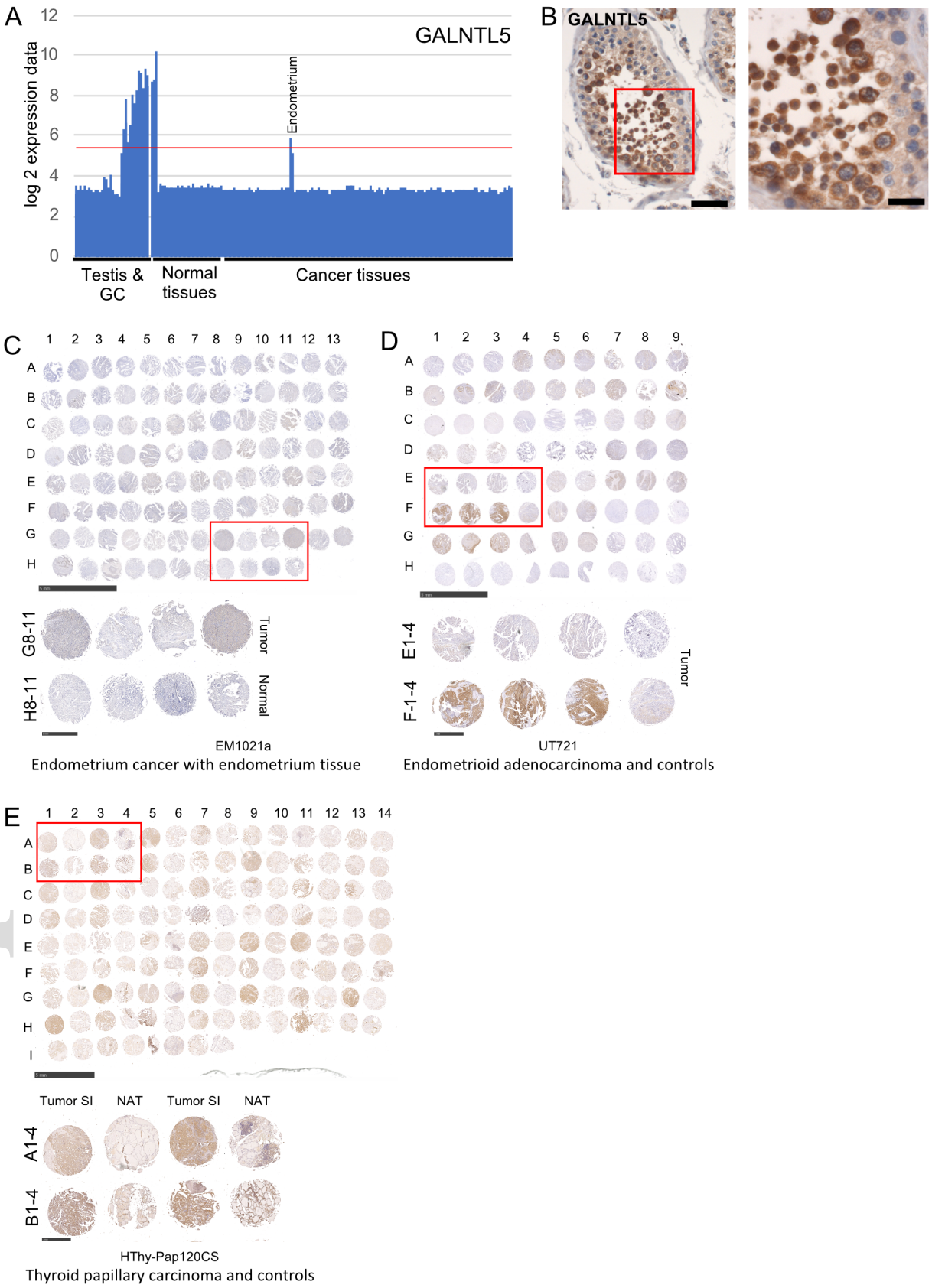
mol2_12900_f6.tiff

Accepted Article



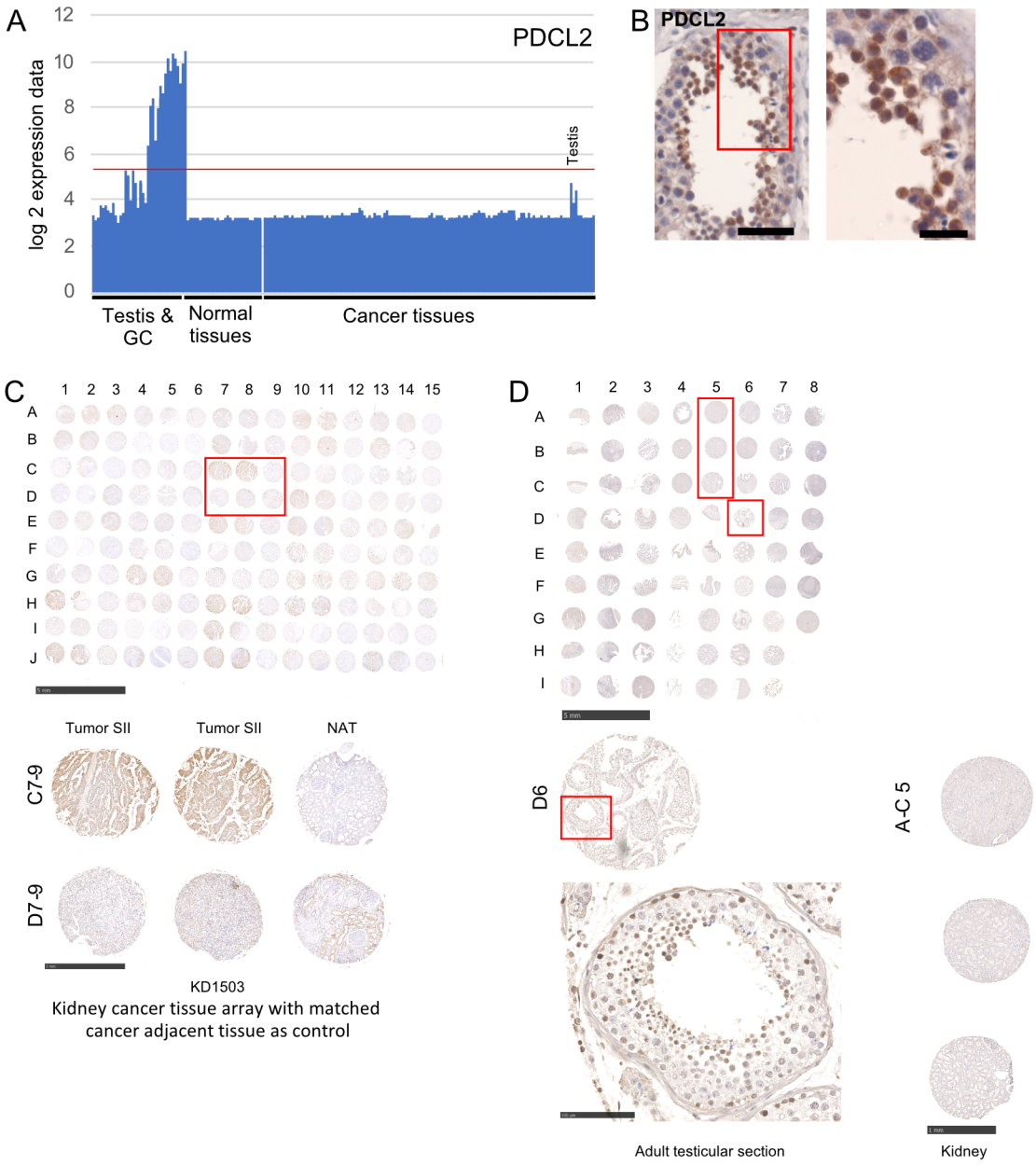
mol2_12900_f7.tiff

Accepted Article



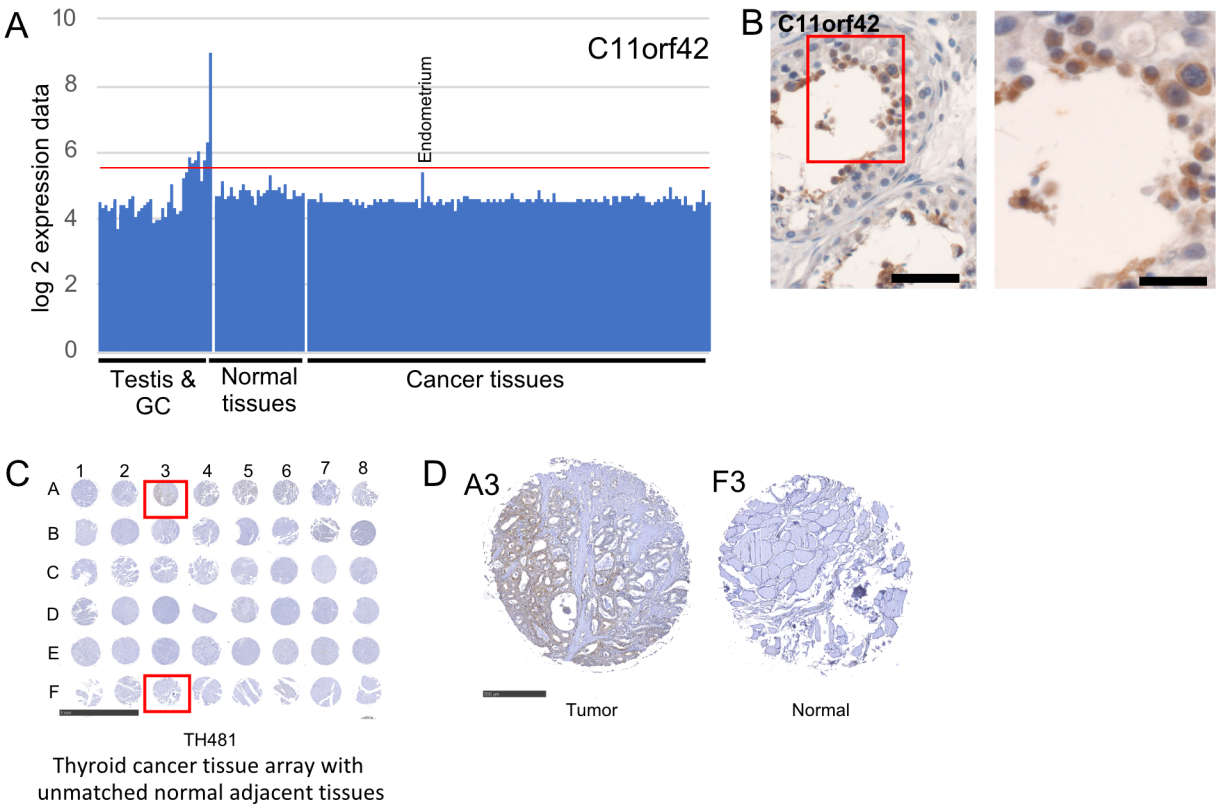
mol2_12900_f8.tiff

Accepted Article



mol2_12900_f9.tiff

Accepted Article



mol2_12900_f10.tiff