THYROID
Volume 27, Number 2, 2017

DOI: 10.1089/thy.2016.0285

Cell Cycle M-Phase Genes Are Highly Upregulated in Anaplastic Thyroid Carcinoma

Paul Weinberger,[1,2,3,*] Sithara Raju Ponny,[1] Hongyan Xu,[4] Shan Bai,[1] Robert Smallridge,[5] John Copland,[6] and Ashok Sharma[1,4]

Background: Anaplastic thyroid carcinoma (ATC) accounts for only 3% of thyroid cancers, yet strikingly, it accounts for almost 40% of thyroid cancer deaths. Currently, no effective therapies exist. In an effort to identify ATC-specific therapeutic targets, we analyzed global gene expression data from multiple studies to identify ATC-specific dysregulated genes.
Methods: The National Center for Biotechnology Information Gene Expression Omnibus database was searched for high-throughput gene expression microarray studies from human ATC tissue along with normal thyroid and/or papillary thyroid cancer (PTC) tissue. Gene expression levels in ATC were compared with normal thyroid or PTC using seven separate comparisons, and an ATC-specific gene set common in all seven comparisons was identified. We investigated these genes for their biological functions and pathways.
Results: There were three studies meeting inclusion criteria, (including 32 ATC patients, 69 PTC, and 75 normal). There were 259 upregulated genes and 286 downregulated genes in ATC with at least two-fold change in all seven comparisons. Using a five-fold filter, 36 genes were upregulated in ATC, while 40 genes were downregulated. Of the 10 top globally upregulated genes in ATC, 4/10 (*MMP1*, *ANLN*, *CEP55*, and *TFPI2*) are known to play a role in ATC progression; however, 6/10 genes (*TMEM158*, *CXCL5*, *E2F7*, *DLGAP5*, *MME*, and *ASPM*) had not been specifically implicated in ATC. Similarly, 3/10 (*SFTA3*, *LMO3*, and *C2orf40*) of the most globally downregulated genes were novel in this context, while 7/10 genes (*SLC26A7*, *TG*, *TSHR*, *DUOX2*, *CDH1*, *PDE8B*, and *FOXE1*) have been previously identified in ATC. We experimentally validated a significant correlation for seven transcription factors (*KLF16*, *SP3*, *ETV6*, *FOXC1*, *SP1*, *EGFR1*, and *MAFK*) with the ATC-specific genes using microarray analysis of ATC cell lines. Ontology clustering of globally altered genes revealed that ''mitotic cell cycle'' is highly enriched in the globally upregulated gene set (44% of top upregulated genes, p-value $<10^{-30}$).
Conclusions: By focusing on globally altered genes, we have identified a set of consistently altered biological processes and pathways in ATC. Our data are consistent with an important role for M-phase cell cycle genes in ATC, and may provide direction for future studies to identify novel therapeutic targets for this disease.

Keywords: anaplastic thyroid cancer, gene expression, translational research, cell cycle, mitotic phase

Introduction

Unlike most other cancers, the incidence of epithelial thyroid cancer is increasing worldwide. The most common cancer arising from thyroid follicular cells is broadly classified under the appellation differentiated thyroid cancer (DTC) and is comprised of papillary thyroid cancer (PTC), follicular thyroid cancer, and Hürthle cell carcinoma. Unlike most cancers, DTC is considered curable and in the vast majority of cases has >90% cure rates. This contrasts starkly with the fourth subtype of thyroid malignancy arising from follicular cells, anaplastic thyroid carcinoma (ATC), which is highly aggressive and has no effective treatment. Although ATC accounts for only 3% of thyroid cancer, it is responsible for up to 40% of thyroid cancer related deaths. In most cases, ATC patients do not survive more than a year post detection (1). ATC is arguably the most lethal of all human malignancies and carries a median survival time of only 4.3

[1]Center for Biotechnology and Genomic Medicine, [2]Department of Otolaryngology, [3]Augusta University Cancer Center, and [4]Department of Biostatistics and Epidemiology, Medical College of Georgia at Augusta University, Augusta, Georgia.
[5]Division of Endocrinology and [6]Department of Cancer Biology, Mayo Clinic, Jacksonville, Florida.
**Present address:* Louisiana State University Health Sciences Center, Shreveport, Louisiana.

months, and the 5-year overall survival rate is less than 5%. Clinically evident distant metastases are seen in more than half of the patients at time of diagnosis. Even those without metastases are considered to have advanced system disease, and all ATC patients are classified by the American Joint Committee on Cancer staging as stage IV at diagnosis. In addition to the high mortality, ATC carries high morbidity in that almost all patients develop airway compromise due to local disease progression in the neck. In light of its rapid growth rate, it is somewhat surprising that ATC is not sensitive to any current systemic therapies, including radiation or chemotherapy. Although several promising targeted therapies (2–5) and multimodal treatment strategies (6,7) have been explored, to date no therapy has been definitively proven to prolong survival in these patients. Clearly, novel treatments are desperately needed.

Large-cohort gene expression studies comparing cancer to control are powerful tools in uncovering novel oncogene pathways; however, they often suffer from lack of reproducibility. Indeed, differential genes expression in any one experiment may be more due to confounding factors than specific cancer effects. In light of this and the stark differences in survival between DTC and ATC, we sought to address this deficiency and find genes and pathways uniquely upregulated only in ATC tumors. In this study, we used gene expression data from three major studies (GSE29265, GSE33630, and GSE65144) comprising 32 ATC, 69 PTC, and 75 normal thyroid patient tissue samples. We compared gene expression profiles of ATC with normal thyroid and PTC using seven separate comparisons, and identified an ATC-specific set of genes altered in all seven comparisons. Differentially expressed genes were further annotated to biological processes, cellular components and biological pathways to identify functional aspects of dysregulated genes. These analyses revealed that cell cycle–related processes, specifically mitotic cell cycle genes, were highly enriched upregulated in ATC. In addition to confirming several known ATC-related genes, our approach of combined analyses of gene expression studies provides a novel set of genes involved in ATC that may be relevant for future therapeutic strategies in this disease.

Materials and Methods

Gene expression data analyses

The National Center for Biotechnology Information Gene Expression Omnibus (GEO) database was searched for high-throughput gene expression microarray studies that included human ATC tissue as well as normal thyroid and/or PTC tissue. Microarray studies based on immortalized ATC cell lines were specifically excluded. The three studies using an identical platform (Affymetrix Human Genome U133 Plus 2.0 Array) were selected for further analyses (GSE29265: 9 ATC, 20 PTC, 20 normal; GSE33630: 11 ATC, 49 PTC, 45 normal; and GSE65144: 12 ATC, 10 normal). Microarray raw expression data (CEL files) were downloaded from the GEO database and each dataset was analyzed separately (Fig. 1). Normalization was performed using the Robust Multiarray Average package, followed by annotation of probes to human genes using the hgu133plus2.db package. We used the Linear Models for Microarray Bioconductor package for differential expression analysis. To discover ATC-specific genes, gene expression levels in the ATC group were compared with PTC and normal groups in each dataset separately, resulting in seven different comparisons. We used the False Discovery Rate method to adjust the p-values for multiple testing, which is Benjamini and Hochberg's method to control the false discovery rate. Gene lists significantly differentially expressed in each comparison (using fold change (FC) >2.0 and adjusted p-value <0.01) were filtered. Reproducible ATC-specific genes differentially expressed in all seven comparisons (259 upregulated and 286 downregulated) were selected for further analyses (Fig. 1). In order to understand the expression patterns of differentially expressed genes, we performed cluster analysis using HPCluster program (8).

Biological function and pathway analyses

Gene ontology enrichment analysis was performed on 545 ATC-specific genes using DAVID (9) to provide an overview of the major biological process and cellular components. Ingenuity pathway analysis (IPA) was used to identify canonical pathways represented by the ATC-specific gene set (10). Network analysis was performed to search for defined molecular interactions between genes (gene products). To identify important transcription factors involved in the ATC-specific gene set, we used MEME and TOMTOM software tools. MEME was used to explore 5 kb upstream of specific genes for the motifs of length (6–20 bp), with stringent threshold of e-value $\leq 1 \times 10^{-10}$. Next the TOMTOM and Jolma 2013 database was used to identify transcription factors which are likely to bind significantly (p-value ≤ 0.01) to their corresponding motifs (11). We used the motif–gene relationship from MEME to look into the relationship between transcription factors (TF) and gene network. For each gene, we identified upstream motifs and probable transcription factors at a p-value cut-off of 0.01. We used a custom Perl script to design a TF-gene network list and Cytoscape 3.2.1 was used to visualize this network.

Expression profiling of ATC and normal thyroid cell lines

Gene expression microarray analysis was performed for four ATC cell lines (THJ11T, THJ16T, THJ21T, THJ29T; see reference Marlow *et al.* (12) for characterization of these cell lines) and 3 normal primary thyroid cells (THJ-45N, THJ-101N, THJ-122N) using Affymetrix Human Genome U133 Plus 2.0 Array. Briefly, RNA was extracted from cell lines using TRIzol (Invitrogen) and chloroform (Sigma). The 18S/28S bands were verified on a 1% agarose gel. RNA quality was assessed by Agilent Bioanalyzer. The RNA products were column-purified (Affymetrix) and then *in vitro* transcribed to generate biotin-labeled cRNA. The *in vitro* transcription products were column purified, fragmented, and hybridized onto Affymetrix U133 Plus 2.0 GeneChips® at 45°C for 16 hours. Subsequent to hybridization, the arrays were washed and stained with streptavidin-phycoerythrin, then scanned in an Affymetrix GeneChip® Scanner 3000. All control parameters were confirmed to be within normal range before normalization and data reduction was initiated. Raw data was processed by MAS.5 (Affymetrix) and analyzed using GeneSpring GX10.

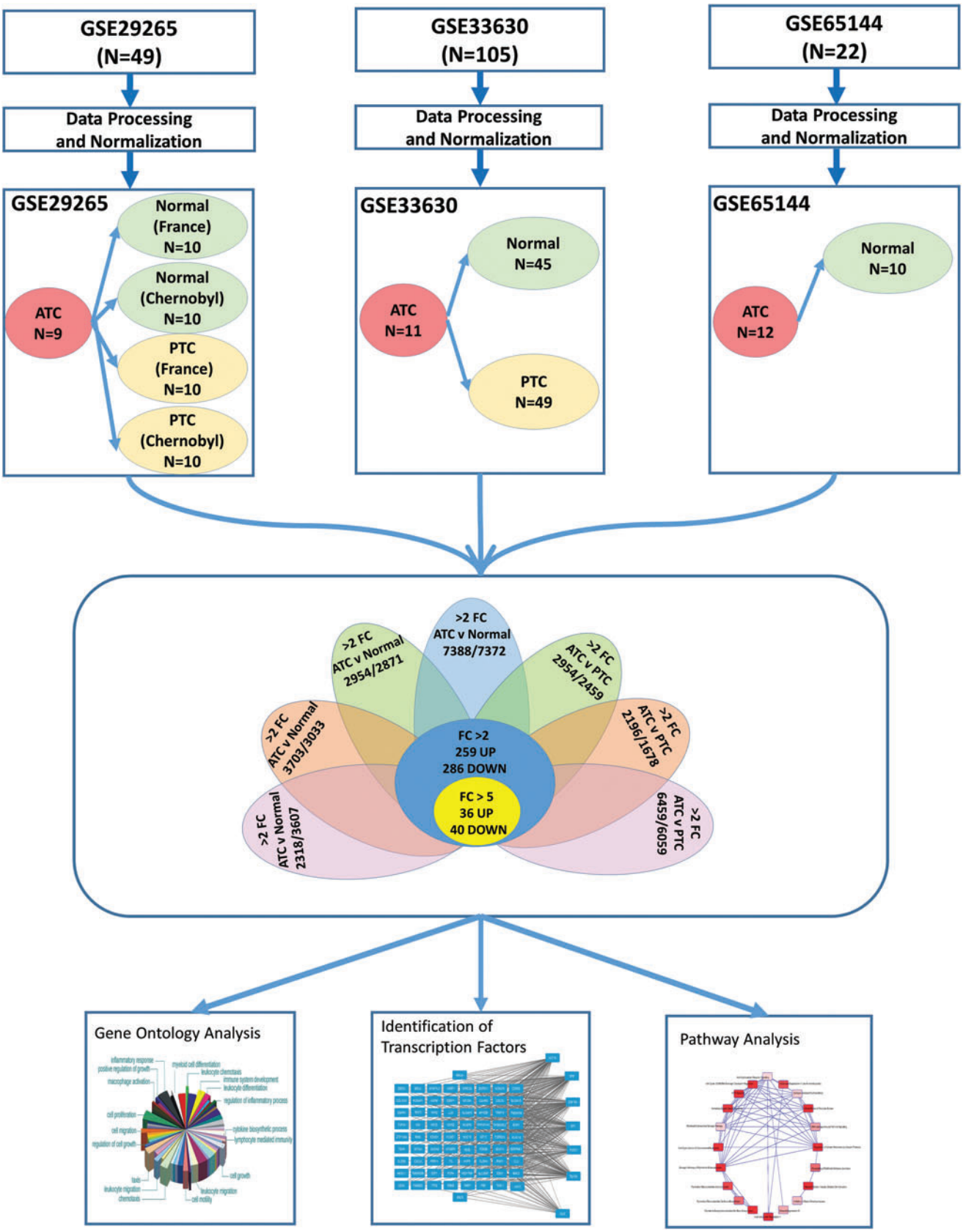

FIG. 1. Flowchart summarizing the study design. Three studies that included human anaplastic thyroid carcinoma (ATC) tissue for gene expression microarrays were selected. Microarray raw expression data (CEL files) were downloaded from the Gene Expression Omnibus database and each dataset was analyzed separately. Seven different comparisons were made to compare ATC group with normal and papillary thyroid cancer (PTC) groups. ATC-specific genes common in all seven comparisons were identified. These genes were further investigated for their biological functions, pathways, and transcription factors. FC, fold change. Color images available online at www.liebertpub.com/thy

Results

Discovery of novel ATC-specific genes

A total of seven comparisons were made to discover genes differentially expressed in ATC as compared with normal or PTC groups (Fig. 1). Enumeration of genes with significant (adjusted $p < 0.01$) changes in each comparison are presented in Table 1. There were 545 genes (259 upregulated and 286 downregulated) in ATC with at least two-fold change globally present in 7/7 (100%) comparisons (Table 1). Using a more stringent five-fold filter, 36 genes were upregulated in ATC (Table 2), while 40 genes were downregulated (Table 3). The heat map representing the expression levels of 76 genes in individual samples is shown in Fig. 2. Unsupervised cluster analysis of these genes revealed three clustering patterns (Fig. 2). Cluster 1 represents a set of genes that have low expression in the majority of ATC samples and high expression in PTC and normal samples. Cluster 2 represents genes with high expression in ATC and low expression in PTC and normal samples. Cluster 3 represents genes with low expression in ATC, high expression in normal, and mixed expression in PTC samples. The 10 most upregulated genes in ATC were *MMP1* (mean FC: 32.6), *TMEM158* (mean FC: 23.1); *CXCL5* (mean FC: 21.3); *ANLN* (mean FC: 17.4); *E2F7* (mean FC: 14.51); *DLGAP5* (mean FC: 13.0); *CEP55* (mean FC: 12.57); *RRM2* (mean FC: 12.5); *TFPI2* (mean FC: 12.4); and *MME* (mean FC: 12.2). Of these top globally upregulated genes, four (*MMP1*, *ANLN*, *CEP55*, and *COL5A1*) have been previously studied in the context of ATC (13–16); however, the remaining six genes have not been specifically implicated in ATC. The ten most downregulated genes in ATC were *SLC26A7* (mean FC: −46.1); *C2orf40* (mean FC: −44.4); *TG* (mean FC: −35.9); *TSHR* (mean FC: −29.9); *SFTA3* (mean FC: −28.2); *DUOX2* (mean FC: −27.6); *CDH1* (mean FC: −25.3); *PDE8B* (mean FC: −22.2); *LMO3* (mean FC: −21.7); and *FOXE1* (mean FC: −20.5). For downregulated genes, only three (*SFTA3*, *LMO3*, and *C2orf40* encoding the ECRG4 protein) were novel in this context, while the seven other genes have been explicitly proposed or studied in the context of ATC tumor biology and reviewed earlier (17).

Mitotic cell cycle is highly enriched in ATC-specific genes

Gene ontology clustering analyses was conducted to associate ATC-specific genes to different biological processes. A total of 35 biological processes were significantly enriched in globally upregulated ATC-specific genes (Table 4). Highly enriched biological processes in the globally upregulated genes are ''M-phase'' (50 genes, $p = 2 \times 10^{-32}$); ''cell cycle phase'' (54 genes, $p = 6 \times 10^{-32}$); ''mitotic cell cycle'' (51 genes, $p = 4 \times 10^{-31}$); ''mitosis'' (41 genes, $p = 2 \times 10^{-29}$); and ''nuclear division'' (41 genes, $p = 2 \times 10^{-29}$). These results are consistent with an important role for M-phase cell cycle genes in ATC. In globally downregulated genes a total of 27 biological processes were significantly enriched (Supplementary Table S1; Supplementary Data are available online at www.liebertpub.com/thy). The most enriched biological processes are ''neuron development'' (16 genes, $p = 5.4 \times 10^{-5}$); ''neuron differentiation'' (18 genes, $p = 8.8 \times 10^{-5}$); ''neuron projection development'' (13 genes, $p = 1.8 \times 10^{-4}$); and ''endocrine system development'' (7 genes, $p = 3.2 \times 10^{-4}$).

Further analysis was conducted to associate ATC-specific genes to different cellular components. A total of 18 cellular components were significantly enriched ($p < 0.05$) in globally upregulated genes in ATC (Table 5). Examination of cellular components revealed that ''spindle'' (27 genes; $p = 8 \times 10^{-18}$); ''microtubule cytoskeleton'' (43 genes, $p = 8 \times 10^{-16}$); and ''chromosome centromeric region'' (21 genes, $p = 2 \times 10^{-12}$) are the most enriched cellular components in the upregulated genes, and interestingly, these cellular components are all involved in the M-phase of cell division. Taken together, these data suggest that cell cycle and specifically M-phase are biological processes critical to ATC. Analysis of downregulated genes revealed a total of 13 cellular components significantly enriched (Supplementary Table S2). The most enriched downregulated cellular components in ATC are

Table 1. Number of Genes with Significant ($p < 0.01$) Changes

Number	*Comparison*	*Upregulated genes*	*Downregulated genes*	*Total*
1	GSE65144 ATC vs. Normal (USA)	2318	3607	5925
2	GSE29265 ATC vs. PTC (Chernobyl)	2954	2459	5413
3	GSE29265 ATC vs. Normal (Chernobyl)	3703	3033	6736
4	GSE29265 ATC vs. PTC (France)	2196	1678	3874
5	GSE29265 ATC vs. Normal (France)	2954	2871	5825
6	GSE33630 ATC vs. Normal (Chernobyl)	7388	7372	14760
7	GSE33630 ATC vs. PTC (Chernobyl)	6459	6059	12518
	More than two-fold in all seven comparisons	259	286	545
	More than five-fold in all seven comparisons	36	40	76

Table 2. Most Overexpressed Genes in Anaplastic Thyroid Carcinoma

Gene symbol	*Gene description*	*Chr location*	*FC1*	*FC2*	*FC3*	*FC4*	*FC5*	*FC6*	*FC7*
MMP1	Matrix metallopeptidase 1 (interstitial collagenase)	chr11q22.3	44.2	23.2	28.1	10.3	32.6	54.3	36.1
TMEM158	Transmembrane protein 158 (gene/pseudogene)	chr3p21.3	34.3	22.5	22.6	16.7	19.9	25.5	20.7
CXCL5	Chemokine (C-X-C motif) ligand 5	chr4q13.3	17.0	15.9	24.9	12.9	24.0	33.8	20.6
ANLN	Anillin, actin binding protein	chr7p15-p14	15.0	15.6	22.2	10.2	19.4	25.2	14.2
E2F7	E2F transcription factor 7	chr12q21.2	12.8	15.2	18.5	11.6	19.7	14.2	9.6
DLGAP5	Discs, large (Drosophila) homolog-associated protein 5	chr14q22.3	8.8	10.9	14.0	8.1	14.3	22.6	12.9
CEP55	Centrosomal protein 55kDa	chr10q23.33	9.1	13.4	16.8	9.2	15.4	14.6	9.5
RRM2	Ribonucleotide reductase M2	chr2p25-p24	8.6	9.1	16.9	7.7	15.0	21.8	8.5
TFPI2	Tissue factor pathway inhibitor 2	chr7q22	14.9	6.7	6.9	5.5	7.4	25.6	20.1
MME	Membrane metallo-endopeptidase	chr3q25.2	14.2	11.9	14.8	10.3	12.0	10.8	11.5
ASPM	asp (abnormal spindle) homolog, microcephaly associated (Drosophila)	chr1q31	9.5	8.2	9.0	6.9	9.0	24.2	14.1
UHRF1	Ubiquitin-like with PHD and ring finger domains 1	chr19p13.3	7.2	9.4	16.0	6.6	16.5	18.4	6.8
MELK	Maternal embryonic leucine zipper kinase	chr9p13.2	7.2	8.4	15.9	5.8	13.9	16.4	9.3
LOC100506303	Uncharacterized LOC100506303	—	9.3	6.3	10.0	6.6	9.8	22.3	10.1
TOP2A	Topoisomerase (DNA) II alpha 170kDa	chr17q21-q22	11.1	7.4	11.4	6.2	15.9	14.7	5.7
SERPINE1	Serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1	chr7q22.1	13.5	7.8	7.4	5.9	8.7	15.1	12.6
PBK	PDZ binding kinase	chr8p21.2	6.6	8.6	10.2	7.7	10.4	16.7	10.0
PRR11	Proline rich 11	chr17q22	5.1	10.4	10.2	10.2	11.2	11.7	9.5
NUF2	NDC80 kinetochore complex component, homolog (S. cerevisiae)	chr1q23.3	6.1	9.0	9.7	7.0	9.3	14.6	12.2
COL12A1	Collagen, type XII, alpha 1	chr6q12-q13	9.9	10.5	9.8	8.3	9.3	11.2	8.9
CDC20	Cell division cycle 20	chr1p34.1	6.7	9.2	10.3	10.6	10.6	10.6	7.4
TRIP13	Thyroid hormone receptor interactor 13	chr5p15.33	5.8	8.5	9.0	7.3	9.9	10.4	8.8
NUSAP1	Nucleolar and spindle associated protein 1	chr15q15.1	6.4	5.2	9.0	5.1	10.4	16.2	5.8
LOX	Lysyl oxidase	chr5q23.2	17.4	5.2	6.7	5.2	6.6	9.1	7.9
BUB1	BUB1 mitotic checkpoint serine/threonine kinase	chr2q14	5.1	6.5	7.6	5.0	7.2	16.2	9.9
FOXD1	Forkhead box D1	chr5q12-q13	5.1	8.9	8.3	8.4	7.0	9.5	10.0
TPX2	Microtubule-associated, homolog (*Xenopus laevis*)	chr20q11.2	5.7	5.4	6.7	6.7	7.5	16.4	8.7
BCAT1	Branched chain amino-acid transaminase 1, cytosolic	chr12p12.1	6.2	8.0	11.6	5.2	9.6	10.1	5.7
CDKN3	Cyclin-dependent kinase inhibitor 3	chr14q22	10.2	5.6	6.3	5.4	6.2	12.7	9.6
KIF20A	Kinesin family member 20A	chr5q31	8.2	5.6	6.9	6.3	9.4	11.9	6.7
KIF15	Kinesin family member 15	chr3p21.31	5.2	7.2	8.3	7.3	9.4	9.4	6.5
KIF23	Kinesin family member 23	chr15q23	7.6	6.2	7.5	5.1	7.1	11.0	8.5
FAM83D	Family with sequence similarity 83, member D	chr20q11.22	6.2	5.2	5.8	5.6	7.0	13.3	9.3
NCAPG	Non-SMC condensin I complex, subunit G	chr4p15.33	6.3	5.9	6.6	5.8	6.7	11.4	9.0
DEPDC1	DEP domain containing 1	chr1p31.2	5.9	6.4	6.5	6.4	6.7	9.0	8.5
LOC375295	Uncharacterized LOC375295	chr2q31.1	9.1	6.9	6.2	5.8	5.9	5.0	6.1

All listed genes had greater than five-fold change in all seven comparisons; False Discovery Rate (FDR) adjusted p-value <0.01.
FC1: GSE65144: ATC ($n=12$) vs. Normal ($n=10$).
FC2: GSE29265: ATC ($n=9$) vs. PTC from Chernobyl Tissue Bank ($n=10$).
FC3: GSE29265: ATC ($n=9$) vs. Normal from Chernobyl Tissue Bank ($n=10$).
FC4: GSE29265: ATC ($n=9$) vs. PTC from the Ambroise Paré Hospital, France ($n=10$).
FC5: GSE29265: ATC ($n=9$) vs. Normal from the Ambroise Paré Hospital, France ($n=10$).
FC6: GSE33630: ATC ($n=11$) vs. Normal from Chernobyl Tissue Bank ($n=45$).
FC7: GSE33630: ATC ($n=11$) vs. PTC from Chernobyl Tissue Bank ($n=49$).
Chr, chromosome; FC, fold change.

''anchoring junction'' (12 genes, $p=4.1\times10^{-5}$); ''endoplasmic reticulum'' (29 genes, $p=2.8\times10^{-4}$); and ''adherens junction'' (10 genes, $p=4.2\times10^{-4}$).

Next, IPA was used to associate the 545 ATC-specific genes (including up- and downregulated) to known canonical pathways. IPA analyses revealed that the 31 pathways most affected by ATC-specific genes include canonical pathways likely to play an important role in cell cycle M-phase: ''cell cycle: G2/M DNA damage checkpoint regulation'' (49 genes, $p=2.0\times10^{-4}$); ''mitotic roles of polo-like kinase'' (66

Table 3. Most Underexpressed Genes in Anaplastic Thyroid Carcinoma

Gene symbol	*Gene description*	*Chr location*	*FC1*	*FC2*	*FC3*	*FC4*	*FC5*	*FC6*	*FC7*
SLC26A7	Solute carrier family 26, member 7	chr8q23	−101.8	−10.7	−23.0	−9.5	−26.1	−116.0	−36.0
C2orf40	Chromosome 2 open reading frame 40	chr2q12.2	−25.7	−37.0	−27.8	−29.2	−25.2	−90.9	−75.3
TG	Thyroglobulin	chr8q24	−45.2	−11.9	−15.6	−9.5	−15.0	−88.5	−65.6
TSHR	Thyrotropin receptor	chr14q31	−40.9	−19.9	−21.7	−11.6	−15.6	−57.8	−41.9
SFTA3	Surfactant associated 3	chr14q13.3	−47.9	−23.6	−15.5	−25.6	−16.7	−29.5	−39.0
DUOX2	Dual oxidase 2	chr15q15.3	−14.7	−24.0	−16.3	−14.8	−12.2	−53.6	−57.8
CDH1	Cadherin 1, type 1, E-cadherin (epithelial)	chr16q22.1	−30.4	−7.0	−6.8	−6.6	−7.4	−65.5	−53.8
PDE8B	Phosphodiesterase 8B	chr5q13.3	−12.0	−9.4	−9.3	−6.1	−8.7	−54.3	−55.9
LMO3	LIM domain only 3 (rhombotin-like 2)	chr12p12.3	−13.1	−23.0	−11.6	−17.1	−8.8	−22.0	−56.7
FOXE1	Forkhead box E1 (thyroid transcription factor 2)	chr9q22	−38.4	−8.5	−10.4	−5.7	−11.1	−43.2	−26.8
ZBED2	Zinc finger, BED-type containing 2	chr3q13.2	−21.1	−17.5	−22.3	−14.5	−19.0	−29.7	−17.9
RMST	Rhabdomyosarcoma 2 associated transcript (non-protein coding)	chr12q23.1	−18.6	−30.0	−19.5	−26.6	−20.7	−11.5	−14.4
TXNL1	Thioredoxin-like 1	chr18q21.31	−17.7	−6.2	−12.5	−6.7	−15.9	−34.9	−13.0
ENPP5	Ectonucleotide pyrophosphatase/ phosphodiesterase 5 (putative)	chr6p21.1	−26.1	−7.9	−5.9	−6.3	−6.3	−27.7	−24.2
TCERG1L	Transcription elongation regulator 1-like	chr10q26.3	−5.5	−16.6	−7.4	−14.2	−7.2	−16.2	−33.3
TMEM30B	Transmembrane protein 30B	chr14q23.1	−16.9	−11.3	−9.2	−10.4	−10.0	−23.2	−19.0
MAL2	Mal, T-cell differentiation protein 2 (gene/pseudogene)	chr8q23	−26.1	−12.1	−8.9	−10.9	−11.6	−15.1	−12.3
PCP4	Purkinje cell protein 4	chr21q22.2	−27.1	−7.6	−13.1	−8.8	−18.0	−13.9	−5.3
PPP1R14C	Protein phosphatase 1, regulatory (inhibitor) subunit 14C	chr6q24.3	−14.0	−10.8	−10.5	−9.6	−16.4	−18.9	−12.1
STXBP6	Syntaxin binding protein 6 (amisyn)	chr14q12	−15.3	−11.6	−5.2	−10.9	−6.1	−17.5	−24.7
SLC6A13	Solute carrier family 6 (neurotransmitter transporter, GABA), member 13	chr12p13.3	−14.3	−7.4	−9.7	−10.2	−14.6	−20.6	−9.4
CD24	CD24 molecule	chr6q21	−21.5	−6.2	−6.3	−6.3	−7.5	−20.6	−16.1
NKX2-1	NK2 homeobox 1	chr14q13	−19.3	−8.0	−7.0	−8.3	−7.3	−15.6	−18.4
CLDN8	Claudin 8	chr21q22.11	−18.3	−7.9	−11.8	−6.0	−15.8	−17.3	−6.2
HSD17B6	Hydroxysteroid (17-beta) dehydrogenase 6	chr12q13	−22.0	−5.4	−11.3	−5.9	−12.9	−18.7	−5.7
ID4	Inhibitor of DNA binding 4, dominant negative helix-loop-helix protein	chr6p22.3	−14.1	−5.6	−9.1	−5.1	−10.9	−22.3	−9.3
KCNJ16	Potassium inwardly-rectifying channel, subfamily J, member 16	chr17q24.3	−8.4	−11.5	−8.2	−7.6	−5.9	−18.1	−16.4
CLU	Clusterin	chr8p21	−6.1	−14.3	−7.3	−7.7	−5.6	−13.1	−21.7
ESRP1	Epithelial splicing regulatory protein 1	chr8q22.1	−10.9	−8.6	−5.8	−6.9	−5.7	−15.7	−17.7
MAOA	Monoamine oxidase A	chrXp11.3	−9.2	−8.8	−6.1	−8.7	−6.8	−14.3	−15.0
NEBL	Nebulette	chr10p12	−12.0	−6.4	−9.0	−5.1	−9.2	−17.9	−7.4
KLHL14	Kelch-like family member 14	chr18q12.1	−9.3	−7.7	−7.3	−5.5	−7.2	−18.4	−11.4
MUC15	Mucin 15, cell surface associated	chr11p14.3	−11.7	−7.6	−5.6	−7.8	−6.2	−14.3	−12.8
HOPX	HOP homeobox	chr4q12	−12.6	−7.9	−6.0	−7.1	−6.1	−13.3	−11.0
ATP13A4	Atpase type 13A4	chr3q29	−7.8	−12.0	−6.3	−10.3	−6.6	−7.4	−12.6
CLDN3	Claudin 3	chr7q11.23	−5.8	−7.1	−7.0	−5.6	−7.1	−14.1	−11.2
AFAP1L2	Actin filament associated protein 1-like 2	chr10q25.3	−6.6	−6.9	−9.7	−5.5	−9.0	−9.6	−5.1
GPRC5A	G protein-coupled receptor, family C, group 5, member A	chr12p13	−8.8	−5.2	−6.1	−5.1	−5.0	−9.4	−8.0
SH3BGRL2	SH3 domain binding glutamic acid-rich protein like 2	chr6q14.1	−5.8	−6.4	−5.7	−7.3	−6.7	−7.9	−7.5
MYO5B	Myosin VB	chr18q21	−7.0	−6.2	−5.0	−5.7	−5.3	−9.0	−9.2

All listed genes had greater than five-fold change in all seven comparisons; FDR adjusted p-value <0.01.
FC1: GSE65144: ATC ($n=12$) vs. Normal ($n=10$).
FC2: GSE29265: ATC ($n=9$) vs. PTC from Chernobyl Tissue Bank ($n=10$).
FC3: GSE29265: ATC ($n=9$) vs. Normal from Chernobyl Tissue Bank ($n=10$).
FC4: GSE29265: ATC ($n=9$) vs. PTC from the Ambroise Paré Hospital, France ($n=10$).
FC5: GSE29265: ATC ($n=9$) vs. Normal from the Ambroise Paré Hospital, France ($n=10$).
FC6: GSE33630: ATC ($n=11$) vs. Normal from Chernobyl Tissue Bank ($n=45$).
FC7: GSE33630: ATC ($n=11$) vs. PTC from Chernobyl Tissue Bank ($n=49$).

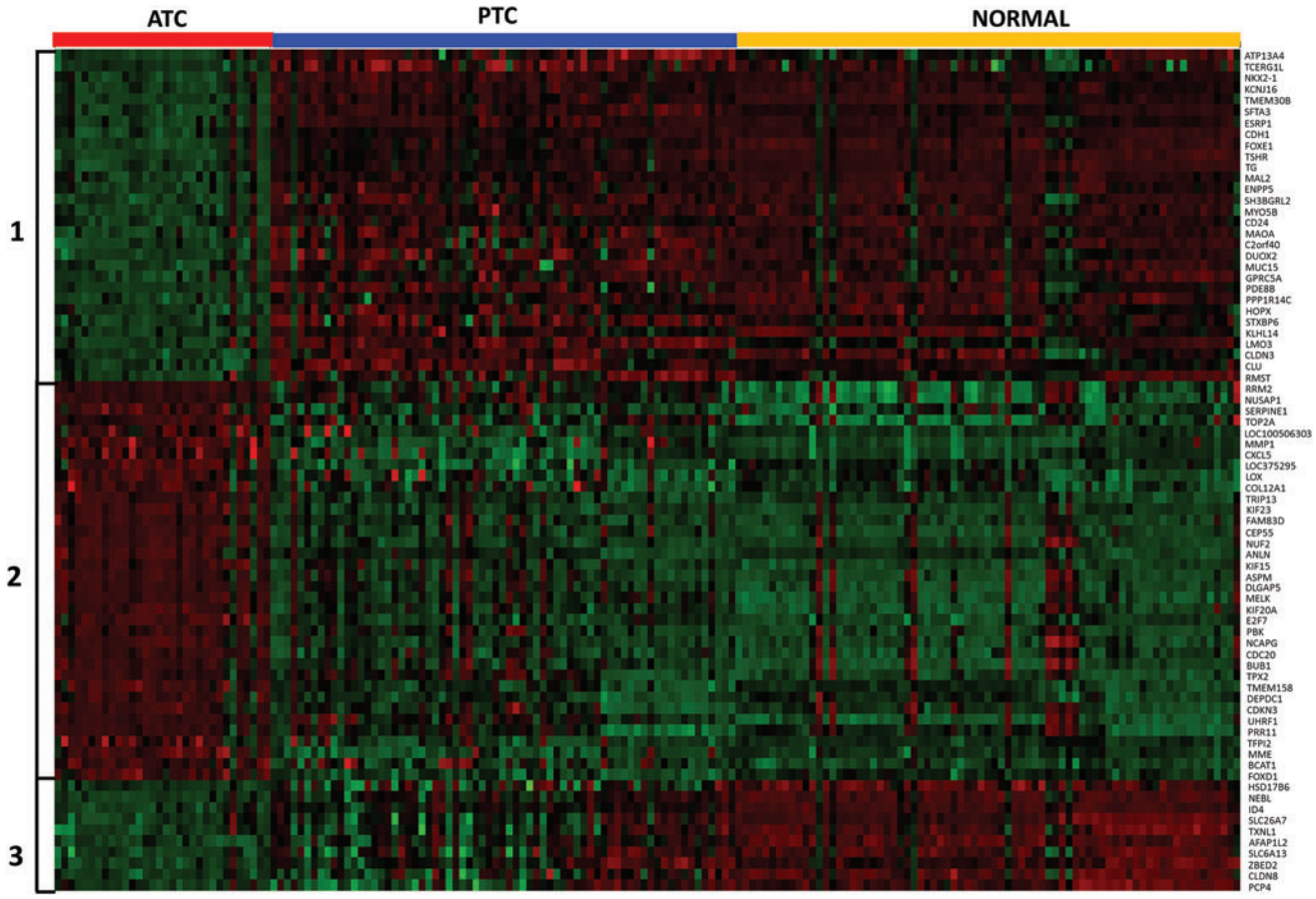

FIG. 2. Heatmap represents clustering of 76 ATC-specific genes into three distinct clusters. Each sample is represented as a column, whereas each gene is represented as a row. Direction and magnitude of expression is represented by the color ranging from high expression (red) to low expression (green). Color images available online at www.liebertpub.com/thy

genes, $p = 2.3 \times 10^{-4}$); ''salvage pathways of pyrimidine ribonucleotides'' (93 genes, $p = 5.4 \times 10^{-4}$); ''pyrimidine deoxyribonucleotides de novo biosynthesis 1'' (22 genes, $p = 2.0 \times 10^{-3}$); ''regulation of cellular mechanics by calpain protease'' (57 genes, $p = 3.0 \times 10^{-3}$); ''ATM signaling'' (59 genes, $p = 3.6 \times 10^{-3}$); and ''cell cycle control of chromosomal replication'' (27 genes, $p = 4.4 \times 10^{-3}$) (Fig. 3). The majority of genes in these pathways showed upregulation. Together, these data reinforce the likelihood that cell cycle M-phase may play a major role in ATC tumor biology.

Network analysis of ATC-specific genes

IPA software was further used to explore functional relationships between these up- and downregulated genes based on known interactions using multiple datasets, resulting in 25 networks. The four highest scoring networks (score >39) are shown in Fig. 4. The top diseases and functions associated with these three networks are cancer, cell death and survival, organismal injury and abnormalities, cell cycle, cellular movement, reproductive system development and function, and cellular assembly and organization. The data on all 25 networks is presented in Supplementary Table S3.

Identification of transcription factors

Transcription factor regulation is a critical dimension of gene expression regulation. We used MEME and TOMTOM software tools to identify transcription factors predicted to affect the expression of the 545 ATC-specific genes (see Methods). The resulting transcription-gene network was visualized using Cytoscape. There were 23 transcription factors predicted to be involved in the regulation of the 545 ATC-specific genes (Fig. 5, p-value <0.005). These 23 transcription factors were enriched for a cluster of transcription factors belonging to the SP/KLF family (*SP1*, *SP3*, *SP8*, *KLF14*, and *KLF16*). Other transcription factors identified in this analysis were *CPEB1*, *EGR1*, *ELK1*, *ETV6*, *GLI2*, *FOXC1*, *FOXG1*, *MAFK*, *MEF2D*, *ONECUT1*, *ONECUT3*, *PAX1*, *PAX9*, *ZBTB49*, *ZIC3*, *ZFP740*, *ZNF524*, and *ZNF740*.

Experimental validation of identified transcription factors

For experimental validation of identified transcription factors, we performed gene expression microarray analysis of four ATC cell lines (THJ11, THJ16, THJ21, THJ29) along with three normal primary cells (THJ45N, THJ101N, THJ122N). Correlations between expression levels of these transcription factors with their regulated ATC-specific genes were checked. We were able to validate a significant correlation for seven transcription factors (*KLF16*, *SP3*, *ETV6*, *FOXC1*, *SP1*, *EGFR1*, and *MAFK*) with the ATC-specific genes (Fig. 6). The cell line gene expression microarray data is submitted to the GEO database (GSE85457).

Table 4. Overexpressed Biological Processes

Biological process	*Count*	*p-Value*
GO:0000279 ~ M phase	50	2.13×10^{-32}
GO:0022403 ~ cell cycle phase	54	6.10×10^{-32}
GO:0000278 ~ mitotic cell cycle	51	4.84×10^{-31}
GO:0007067 ~ mitosis	41	2.66×10^{-29}
GO:0000280 ~ nuclear division	41	2.66×10^{-29}
GO:0000087 ~ M phase of mitotic cell cycle	41	5.60×10^{-29}
GO:0048285 ~ organelle fission	41	1.39×10^{-28}
GO:0022402 ~ cell cycle process	57	4.91×10^{-28}
GO:0007049 ~ cell cycle	64	3.98×10^{-27}
GO:0051301 ~ cell division	37	7.66×10^{-20}
GO:0007059 ~ chromosome segregation	16	1.20×10^{-09}
GO:0007346 ~ regulation of mitotic cell cycle	20	1.54×10^{-09}
GO:0051726 ~ regulation of cell cycle	27	5.18×10^{-09}
GO:0007017 ~ microtubule-based process	23	3.75×10^{-08}
GO:0000075 ~ cell cycle checkpoint	15	9.89×10^{-08}
GO:0007093 ~ mitotic cell cycle checkpoint	10	1.71×10^{-05}
GO:0010564 ~ regulation of cell cycle process	14	2.13×10^{-05}
GO:0051783 ~ regulation of nuclear division	10	1.95×10^{-04}
GO:0007088 ~ regulation of mitosis	10	1.95×10^{-04}
GO:0007051 ~ spindle organization	9	4.45×10^{-04}
GO:0000226 ~ microtubule cytoskeleton organization	14	4.46×10^{-04}
GO:0051329 ~ interphase of mitotic cell cycle	12	5.68×10^{-04}
GO:0051325 ~ interphase	12	7.62×10^{-04}
GO:0031577 ~ spindle checkpoint	6	8.14×10^{-04}
GO:0030071 ~ regulation of mitotic metaphase/anaphase transition	7	9.88×10^{-04}
GO:0007052 ~ mitotic spindle organization	6	0.0030
GO:0007010 ~ cytoskeleton organization	22	0.0032
GO:0008283 ~ cell proliferation	22	0.0033
GO:0000070 ~ mitotic sister chromatid segregation	7	0.0216
GO:0045841 ~ negative regulation of mitotic metaphase/anaphase transition	5	0.0234
GO:0007094 ~ mitotic cell cycle spindle assembly checkpoint	5	0.0234
GO:0000819 ~ sister chromatid segregation	7	0.0255
GO:0051784 ~ negative regulation of nuclear division	5	0.0347
GO:0045839 ~ negative regulation of mitosis	5	0.0347
GO:0006260 ~ DNA replication	13	0.0426

Table 5. Overexpressed Cellular Components

Cellular components	*Count*	*p-Value*
GO:0005819 ~ spindle	27	8.14×10^{-18}
GO:0015630 ~ microtubule cytoskeleton	43	8.17×10^{-16}
GO:0000775 ~ chromosome, centromeric region	21	2.24×10^{-12}
GO:0000793 ~ condensed chromosome	21	4.92×10^{-12}
GO:0000779 ~ condensed chromosome, centromeric region	16	4.14×10^{-11}
GO:0000777 ~ condensed chromosome kinetochore	15	1.24×10^{-10}
GO:0000776 ~ kinetochore	15	8.00×10^{-09}
GO:0044430 ~ cytoskeletal part	43	1.61×10^{-07}
GO:0005694 ~ chromosome	28	1.35×10^{-06}
GO:0005856 ~ cytoskeleton	50	6.12×10^{-06}
GO:0044427 ~ chromosomal part	24	1.96×10^{-05}
GO:0005874 ~ microtubule	20	3.48×10^{-05}
GO:0043228 ~ non-membrane-bounded organelle	72	5.68×10^{-05}
GO:0043232 ~ intracellular non-membrane-bounded organelle	72	5.68×10^{-05}
GO:0000922 ~ spindle pole	8	0.0007
GO:0005876 ~ spindle microtubule	7	0.0046
GO:0030496 ~ midbody	6	0.0065
GO:0005815 ~ microtubule organizing center	15	0.0336

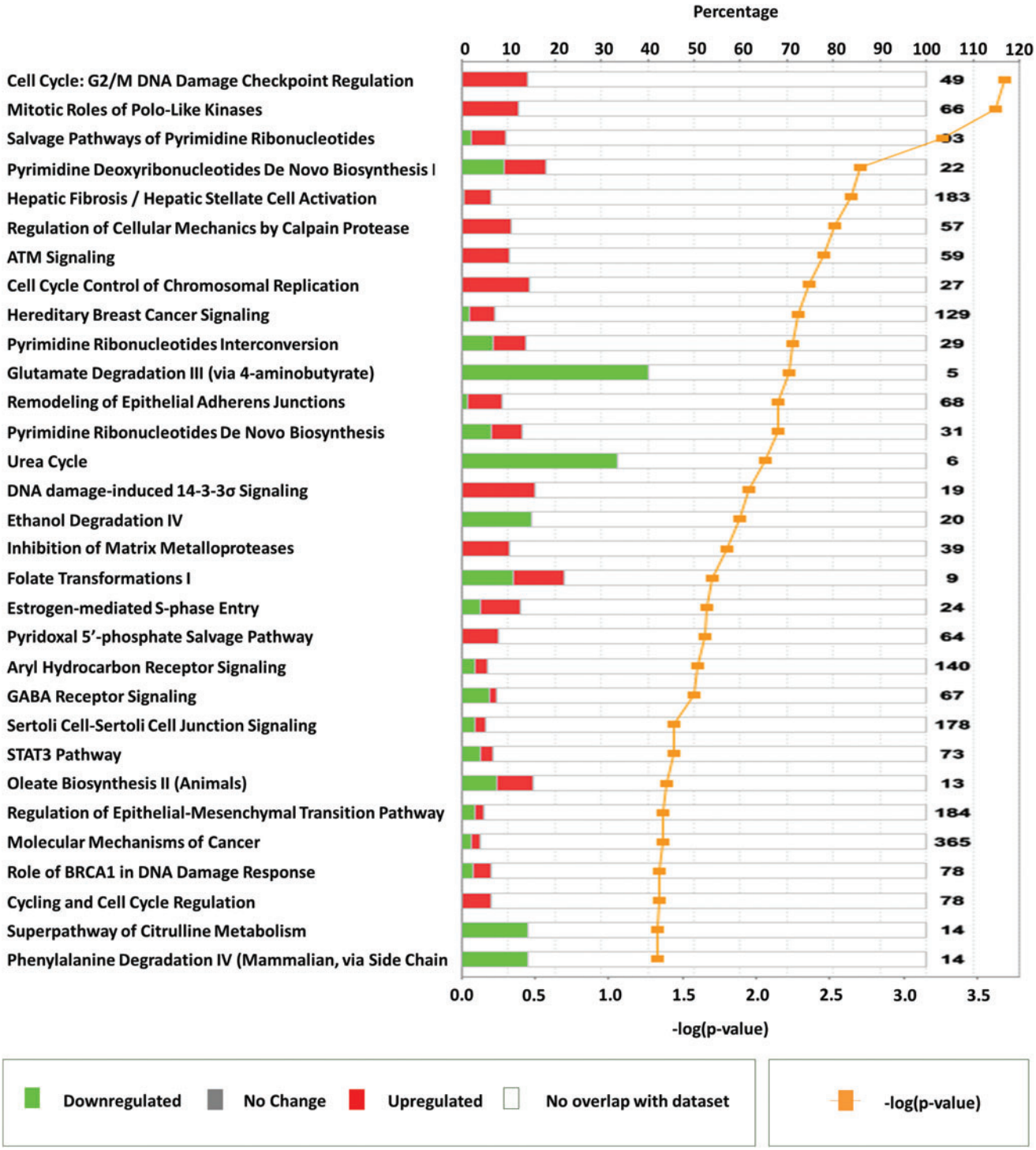

FIG. 3. Important pathways represented by anaplastic thyroid carcinoma–specific genes identified using Ingenuity Pathway Analysis. Horizontal bars represent percentage of overlapping genes in a pathway. Red color signifies upregulation, green color signifies downregulation, and white color represents no overlap with dataset. Yellow line plots the –log (*p*-value) of pathway membership of genes in a specific pathway. Color images available online at www.liebertpub.com/thy

Changes in expression of genes associated with immune infiltration

Recent studies show that immune reaction, involving tumor-associated macrophages and lymphocytic infiltration, is associated with decreased survival in advanced thyroid cancer (18,19). Furthermore, ATC tumors as a group tend to have significant macrophage infiltration, up to 50% of tumor mass in some cases (20). To place our data in context with these findings, we compiled a list of genes associated with ''tumor associated macrophages,'' ''lymphocytic infiltration,'' and ''phagocytosis by macrophages'' (Table 6). We then analyzed

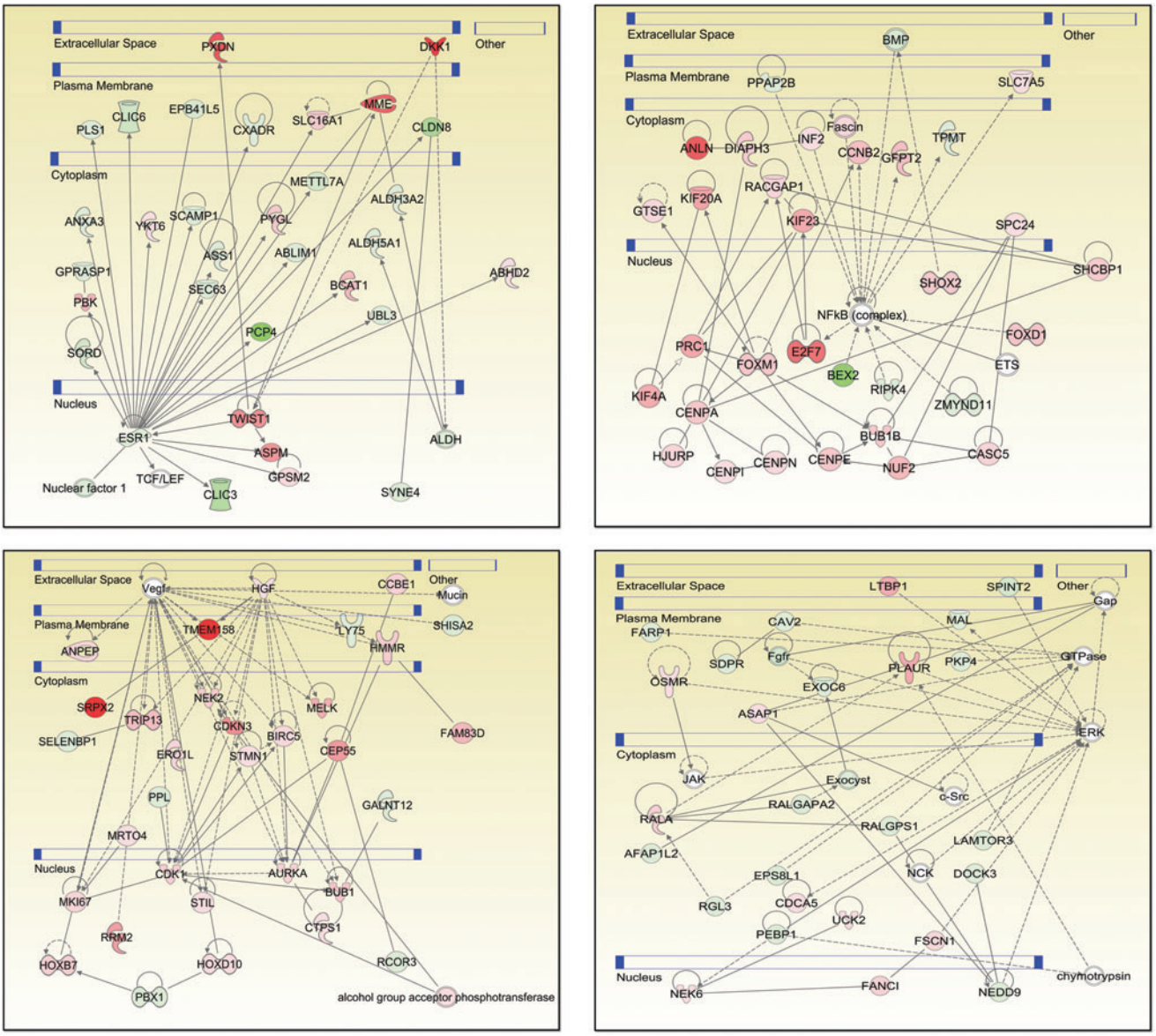

FIG. 4. Four top-scoring networks from Ingenuity Pathway Analysis of anaplastic thyroid carcinoma specific genes. Each gene is represented as a node, and an edge represents an interaction between two nodes. Red nodes indicate upregulated genes, whereas green indicates downregulated genes. White nodes indicate genes not present in ATC-specific gene set. A solid line represents direct functional interaction, while a dotted line represents an indirect interaction. An arrow indicates action of a gene product on a target. Blue bars denote cellular compartments. Color images available online at www.liebertpub.com/thy

the expression of these genes in our ATC samples. We found 1 gene significantly downregulated (FC <0.5-fold) and 27 genes significantly upregulated (FC >2-fold) in ATC samples (Fig. 7). *CSF1R* (mean FC: 3.6) encodes for the cell surface receptor for colony stimulating factor 1 (*CSF1*), a cytokine which controls the production, differentiation, and function of macrophages. Other immune related genes highly upregulated in ATC are *CXCL5* (mean FC: 21.3); *SERPINE1* (mean FC: 10.1); *THBS1* (mean FC: 6.9); and *GJB2* (mean FC: 5.7); *LGALS1* (mean FC: 5.4); *CCR1* (mean FC: 5.2); *HGF* (mean FC: 3.7); and *CLIC4* (mean FC: 3.0). Several members of a family of immunoglobulin Fc receptor genes including *FCGR2A* (mean FC: 5.8); *FCER1G* (mean FC: 4.8); and *FCGR2B* (mean FC: 3.1) were also upregulated in ATC (Table 6).

Discussion

Our study was designed to provide new insights into ATC-specific tumor biology in humans. By leveraging publicly available gene expression data to identify genes altered uniquely and reproducibly in ATC, we provide evidence that cell cycle control genes are dysregulated in ATC. Several previous ATC studies have attempted to leverage global gene expression assays such as expression microarrays to investigate gene dysregulation in ATC. Three of these provided the raw data for the present study. In 2012, Hebrant *et al.* (data deposited online as GSE33630) compared mRNA expression profiles of 11 ATC, 49 PTC tumors, and 45 adjacent normal thyroids and found that the majority of genes altered in ATC

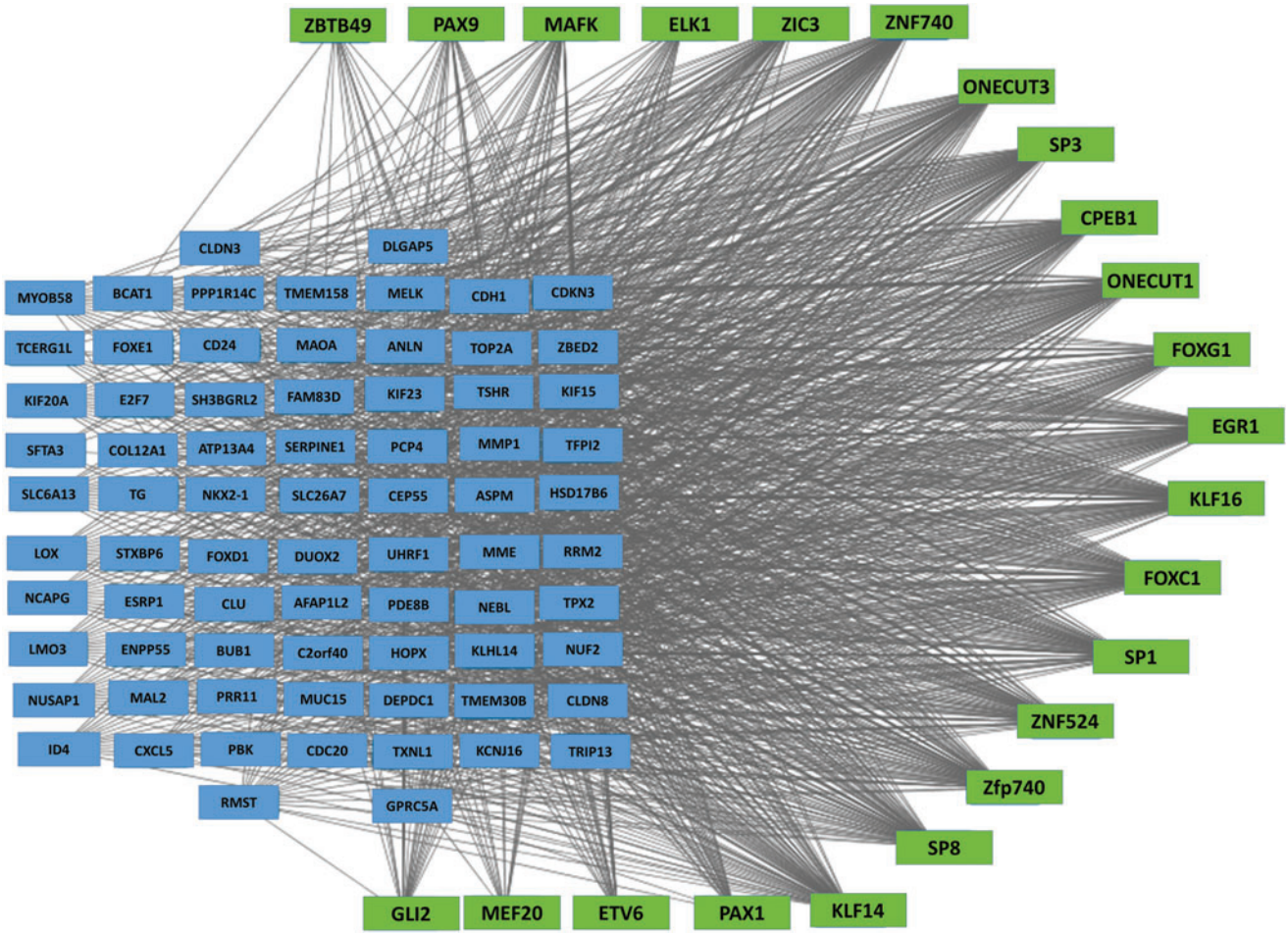

FIG. 5. Transcription factor regulation network of ATC specific genes. Genes (blue rectangles) with at least two-fold change in expression were included in MEME and TOMTOM analysis to identify transcriptional factors (green rectangle) that regulate ATC specific genes. Color images available online at www.liebertpub.com/thy

were also altered in PTC (16). They identified a nine-gene signature that clustered ATC from PTC, consisting of down-regulation of *NELL2*, *SPINT2*, *MARVELED2*, *DUOXA1*, *RPH3AL*, *TBX3*, *PCYOX1*, *C5orf41*, and *PKP4* in ATC. These were subsequently confirmed using RT-PCR. Tomas *et al.* compared mRNA expression in 9 ATCs, 20 PTCs, and 20 normal thyroids in an effort to identify Chernobyl-specific gene expression signatures (data deposited online as GSE29265), but there is no corresponding manuscript at this time. Von Roemeling *et al.* (data deposited online as GSE65144) examined mRNA expression signatures from 12 ATC and 13 normal thyroids (21). They performed extensive validation directed toward alterations in fatty acid metabolism identified by the gene expression study but did not perform gene ontology or pathway clustering. There are also additional genomic or transcriptomic ATC studies leveraging different technologies that have been reported in the literature (22–29). These were not included, as our methods necessitated using data derived from a single platform. It is, however, notable that to date none of these individual studies have led to a therapy for ATC that has demonstrated improved survival in clinical trials.

Therefore, in an effort to identify more robust ATC-specific therapeutic targets, we analyzed data from multiple gene expression microarray studies to identify globally altered genes. Combining microarray gene expression data from multiple laboratories or array platforms can have confounding batch effects leading to false discoveries (30). Many methods exist for removing batch effects from data, however, batch adjustments may bias the results and systematically induce incorrect group differences in downstream analyses (31). Rather than follow a more traditional approach (combined analysis), we instead performed separate comparisons in each dataset followed by identification of statistically significant, differentially expressed genes, which repeatedly appeared in each comparison. We identified a gene signature comprising 259 upregulated and 286 downregulated genes with at least two-fold change globally present in all seven comparisons. This gene set was able to differentiate ATC samples from PTC and normal samples, and could have possible utility as a molecular diagnostic tool in discerning ATC from poorly differentiated PTC, although this is merely a conjecture at this point.

Our data are consistent with an important role for control of the M-phase of the cell cycle in ATC tumor biology, in that a surprising concentration of ATC-specific genes impinge on this pathway. To our knowledge, this is the first report that reveals dysregulation of this pathway in ATC. This is

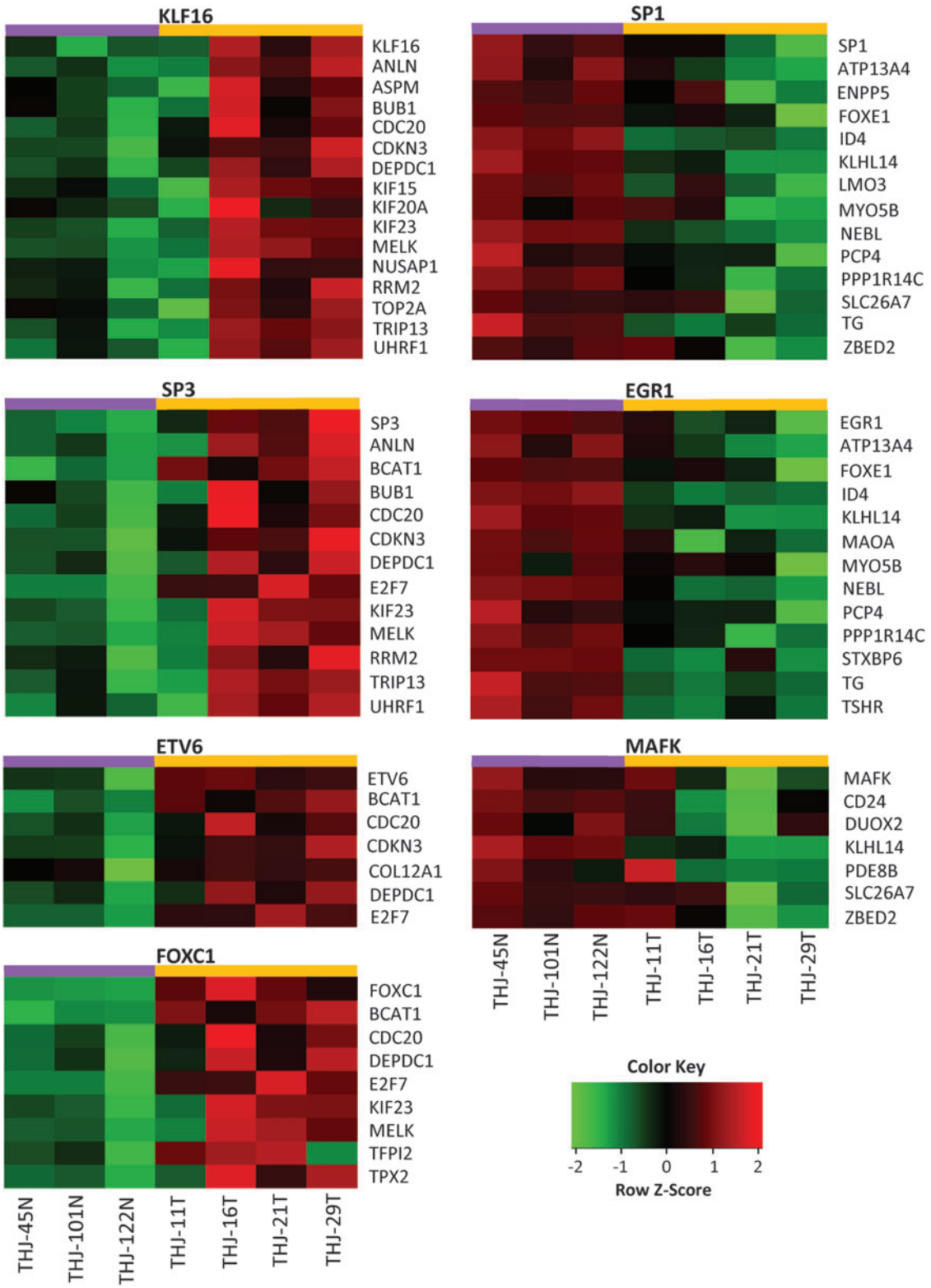

FIG. 6. Experimental validation of identified transcription factors. Gene expression microarray analysis was performed for four ATC cell lines (THJ11, 16, 21, and 29 T cells) and 3 normal primary Thyroid cell lines (THJ45, 101, and 122 N) using Affymetrix Human Genome U133 Plus 2.0 Array. Each row represents one gene, and each column represents one cell line. Red indicates higher expression and green indicates lower expression. Color images available online at www.liebertpub.com/thy

Table 6. Changes in Expression of Genes Associated with Immune Infiltration

Gene symbol	*Gene description*	*Chr location*	*FC1*	*FC2*	*FC3*	*FC4*	*FC5*	*FC6*	*FC7*
MMP1	Matrix metallopeptidase 1 (interstitial collagenase)	chr11q22.3	44.2	23.2	28.1	10.3	32.6	54.3	36.1
CXCL5	Chemokine (C-X-C motif) ligand 5	chr4q13.3	17.0	15.9	24.9	12.9	24.0	33.8	20.6
SERPINE1	Serpin peptidase inhibitor, clade E, member 1	chr7q22.1	13.5	7.8	7.4	5.9	8.7	15.1	12.6
COL1A2	Collagen, type I, alpha 2	chr7q22.1	9.0	3.9	5.2	3.9	5.3	21.5	7.8
THBS1	Thrombospondin 1	chr15q15	5.1	4.5	6.6	3.9	9.8	14.5	4.2
MMP9	Matrix metallopeptidase 9	chr20q 11.2-q13.1	7.0	4.0	6.5	4.1	5.4	14.7	6.6
FCGR2A	Fc fragment of IgG, low affinity IIa, receptor (CD32)	chr1q23	8.1	2.0	3.9	2.8	4.5	12.6	6.6
GJB2	Gap junction protein, beta 2, 26 kDa	chr13q11-q12	8.4	5.0	6.1	3.7	5.6	5.4	5.5
ANPEP	Alanyl (membrane) aminopeptidase	chr15q25-q26	3.9	3.8	3.7	4.2	4.6	11.8	7.2
LGALS1	Lectin, galactoside-binding, soluble, 1	chr22q13.1	8.0	2.9	5.3	4.9	7	7.2	2.3
CCR1	Chemokine (C-C motif) receptor 1	chr3p21	7.1	3.6	5.6	3.9	5.4	6.7	3.8
FCER1G	Fc fragment of IgE, receptor for; gamma polypeptide	chr1q23	7.0	2.2	3.4	2.2	3.7	10.9	4.4
EIF4EBP1	Eukaryotic translation initiation factor 4E binding protein 1	chr8p12	4.6	3.6	3.8	3.6	4.2	6.5	4.9
SPHK1	Sphingosine kinase 1	chr17q25.2	4.9	3.6	5.0	2.9	5.1	5.1	4.0
ACTN1	Actinin, alpha 1	chr14q24| 14q22-q24	5.6	2.6	4.3	2.7	4.6	5.7	2.8
HGF	Hepatocyte growth factor (hepapoietin A; scatter factor)	chr7q21.1	2.5	3.9	4.4	4.5	4.6	3.1	2.8
MAP4K4	Mitogen-activated protein kinase kinase kinase kinase 4	chr2q11.2-q12	3.8	3.2	5.1	2.5	4.1	4.5	2.3
SOD2	Superoxide dismutase 2, mitochondrial	chr6q25.3	3.8	3.6	4.8	2.4	2.9	4.2	3.5
CSF1R	Colony stimulating factor 1 receptor	chr5q32	3.2	2.7	3.6	2.8	4.0	5.2	3.3
C5AR1	Complement component 5a receptor 1	chr19q 13.3-q13.4	2.8	2.3	2.8	2.2	3.0	6.7	4.4
MICB	MHC class 1 polypeptide-related sequence B	chr6p21.3	3.0	2.5	3.2	2.5	3.4	5.0	2.9
CR1	Complement component (3b/4b) receptor	chr1q32	3.8	2.8	2.5	2.7	2.6	3.9	3.9
MMP14	Matrix metallopeptidase 14 (membrane-inserted)	chr14q11-q12	7.8	2.0	2.0	2.3	2.7	2.9	2.1
FCGR2B	Fc fragment of IgG, low affinity IIb, receptor (CD32)	chr1q23	2.7	2.2	2.5	2.3	2.3	5.3	4.3
CLIC4	Chloride intracellular channel 4	chr1p36.11	4.9	2.0	2.3	2.0	2.2	4.1	3.3
SIRPA	Signal-regulatory protein alpha	chr20p13	3.9	2.6	2.9	2.3	2.5	2.8	2.3
TLR4	Toll-like receptor 4	chr9q33.1	2.9	2.4	2.6	2.3	2.1	2.0	2.1
OCLN	Occludin	chr5q13.1	−7.4	−5.5	−4.0	−4.6	−4.5	−9.9	−8.3

All listed genes had greater than two-fold change in all seven comparisons; FDR adjusted p-value <0.05.
FC1: GSE65144: ATC ($n=12$) vs. Normal ($n=10$).
FC2: GSE29265: ATC ($n=9$) vs. PTC from Chernobyl Tissue Bank ($n=10$).
FC3: GSE29265: ATC ($n=9$) vs. Normal from Chernobyl Tissue Bank ($n=10$).
FC4: GSE29265: ATC ($n=9$) vs. PTC from the Ambroise Paré Hospital, France ($n=10$).
FC5: GSE29265: ATC ($n=9$) vs. Normal from the Ambroise Paré Hospital, France ($n=10$).
FC6: GSE33630: ATC ($n=11$) vs. Normal from Chernobyl Tissue Bank ($n=45$).
FC7: GSE33630: ATC ($n=11$) vs. PTC from Chernobyl Tissue Bank ($n=49$).
Ig, immunoglobin; MHC, major histocompatibility complex.

particularly interesting in light of the fact that our analysis made use of publicly available datasets; these were deposited after analysis and reporting of the individual data. As such, the individual up- and downregulated genes were, by definition, altered in these original datasets. In many cases, however, they were ''buried'' in the middle of the data as altered but not in the ''top'' list either by p-value or fold-change value. Our analysis focused attention from these several thousand potentially important genes identified in each individual comparison to a manageable and enriched set of ~500 (FC >2) and ~70 (FC >5) genes. This supports our hypothesis that our combinatorial approach may help to focus attention toward genes that are reproducibly and consistently altered, and thus may be imputed to possibly hold an increased likelihood of playing an important functional role.

Internal validation of our approach is provided in that among the top 10 over- and under-expressed ATC-specific genes, 40% and 70% respectively, have been the focus of previous investigations in ATC. For example, thyroglobulin (TG) and thyrotropin receptor (TSH-R) are universally lost in ATC (17).

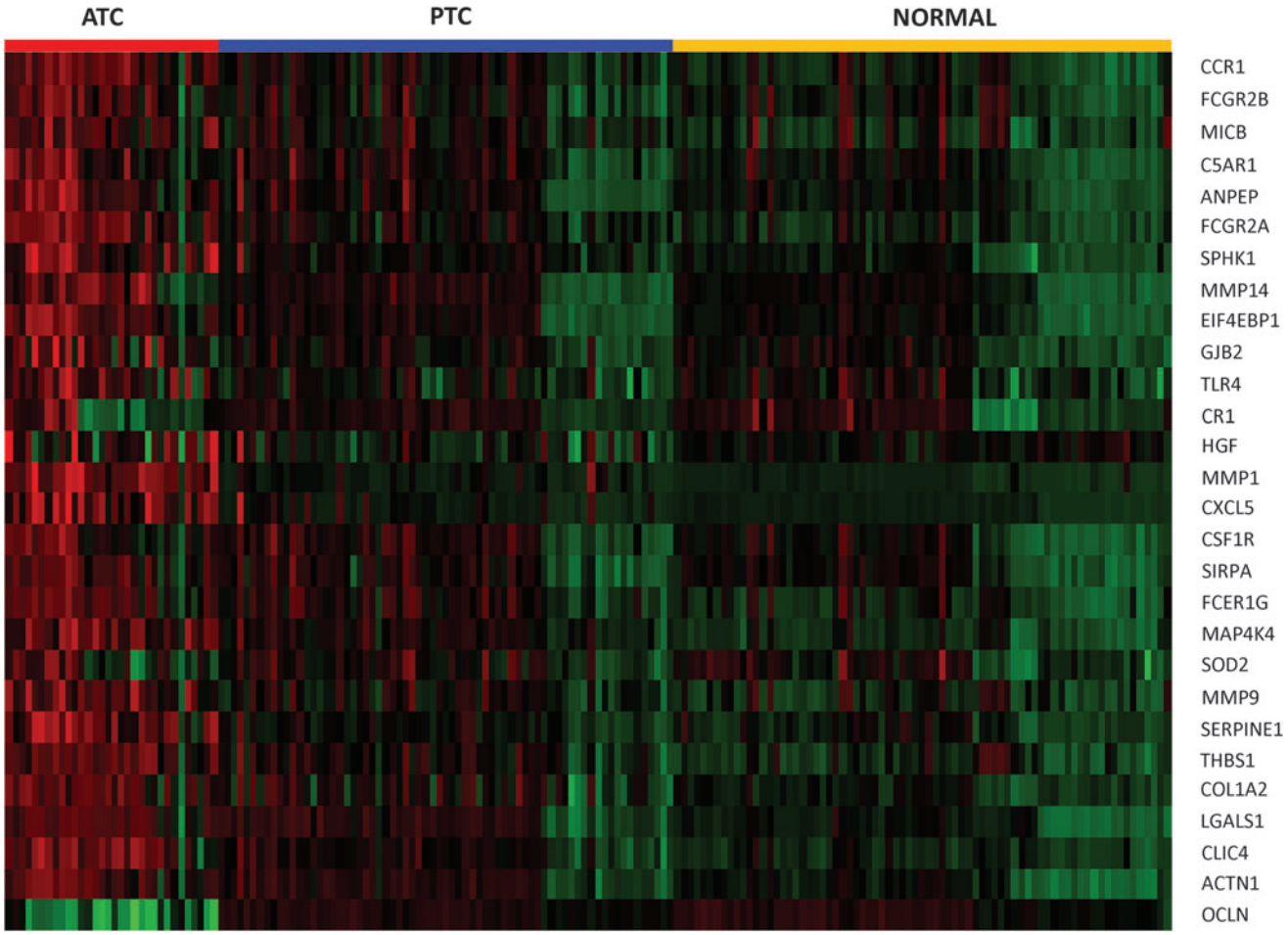

FIG. 7. Heatmap representing 28 immune related genes significantly changed in ATC. Each sample is represented as a column whereas each gene is represented as a row. Direction and magnitude of expression is represented by the color ranging from high expression (red) to low expression (green). Color images available online at www.liebertpub.com/thy

CDH1, an important regulator of epithelial-mesenchymal transition, has likewise been shown to be downregulated in ATC (32,33). Conversely, matrix metalloproteinase-1 (MMP1), important for allowing invasion and metastasis, is highly overexpressed in ATC (16,34). Poorly differentiated thyroid carcinoma specimens have profound deregulation of genes involved in cell adhesion and intracellular junctions, with changes consistent with an epithelial–mesenchymal transition (35). A significant upregulation of the ''cell cycle progression'' was found by the functional profiling of undifferentiated and well-differentiated thyroid tumors (33). These confirmatory findings notwithstanding, the more interesting genes are the ones not yet described in the context of ATC tumor biology, as these may suggest novel therapeutic targets or strategies for this disease. A few of the more promising leads will be further explored here, although space limitations preclude an exhaustive analysis.

One highly upregulated gene, novel in the context of ATC, was C-X-C motif chemokine 5 (CXCL5). CXCL5 is known to play a role in macrophage regulation, and ATC tumors tend to have significant macrophage infiltration. As our study utilized publicly available datasets, determination of macrophage involvement in the ATC tumors included in this study is not possible. It is likely that the *CXCL5* upregulation may reflect inclusion of macrophage RNA in the analyses, but it is also possible that CXCL5 also may play a role in ATC tumor biology. CXCL5 has recently been implicated in mediating cell proliferation, migration and invasion in colorectal cancer and is associated with metastasis and worse prognosis in gastric and breast cancer (36,37). Another novel, ATC-specific overexpressed gene encodes the disks large-associated protein 5 (DLGAP5), a component of the kinetochore responsible for stabilizing microtubules and the spindle apparatus. This protein is relatively unstudied in tumor biology, but recently was shown to have a potential role in hepatocellular cancer. Liao *et al.* demonstrated overexpression of *DLGAP5* mRNA to be common in tumor samples from liver cancer patients, and that *in vitro* RNA-interference mediated silencing of DLGAP5 inhibited proliferation and invasion (38). One of the novel, ATC-specific underexpressed genes was *C2orf40*, which encodes the esophageal cancer–related gene 4 (ECRG4) protein, a tumor suppressor affecting cancer cell migration, invasion and cell cycle regulation and downregulated in esophageal and breast cancer (39,40).

Our transcription factor analysis identified several transcription factors likely to be important in ATC tumor biology. We performed further experimental validation of the identified transcription factors using four ATC cell lines along with

three normal primary cells. We were able to validate a significant correlation for seven transcription factors (*KLF16*, *SP3*, *ETV6*, *FOXC1*, *SP1*, *EGFR1*, and *MAFK*) with the ATC-specific genes (Fig. 6). Specificity-Protein/Krueppel Like Factor (SP/KLF) transcription factors (including *SP1*, *SP3*, and *KLF16* identified in the present study) comprise an emerging group of proteins that are thought to regulate fundamental cellular processes including cell cycle and growth control, metabolic pathways, and apoptosis. There is emerging evidence that expression of genes known to play pivotal roles in metastasis and cell proliferation are regulated by the Sp family of proteins (41). *NTRK3/ETV6* is a known fusion oncogene (42,43) and results in more extensive disease and aggressive pathology of PTC in the pediatric population (44). *FOXC1* is a member of the Forkhead box family transcription factors and is known to promote melanoma by activating the MST1R/PI3K/AKT pathway (45). Although transcription factors have traditionally been considered undruggable targets, recent successes in this arena call this into question (46–49). Taken together, these data suggest that pharmacologic targeting of transcription factors upregulated in ATC (*KLF16*, *SP3*, *ETV6*, and *FOXC1*) might be an area for future investigations in developing novel ATC therapeutics.

Thyroid cancers are heavily infiltrated with macrophages (18) and the density of tumor-associated macrophages is increased in advanced thyroid cancers including ATC (19). The presence of a high density of macrophages may influence the overall gene expression profile in ATC. Several upregulated genes encoding cytokine, chemokine and matrix metalloproteinases are highly expressed by tissue resident macrophages. We found several genes associated with ''tumor-associated macrophages (TAMs),'' ''lymphocytic infiltration'' and ''phagocytosis by macrophages'' including *CSF1R* (mean FC: 3.6), highly upregulated in ATC samples (Table 6). Macrophages depend on CSF-1 for differentiation and survival. The CSF1R inhibitor has been used to target TAMs in a mouse proneural GBM model, which significantly increased survival and regressed established tumors (50). TAMs are emerging as a promising therapeutic target and our results suggest that CSF1R inhibition could have translational potential for ATC treatment.

In conclusion, we have demonstrated that multiple cell cycle M-phase genes are highly upregulated in ATC. Several of the genes identified as ATC-specific are novel in the context of ATC tumor biology and provide a strong rationale supporting their potential as possible therapeutic targets in ATC. Additionally, transcription factor regulation of ATC-specific gene sets suggest the SP/KLF transcription factor family as additional potential therapeutic targets. Our data strongly suggest that therapeutic strategies targeting processes critical for cell cycle mitosis may be of particular value in ATC and deserve further investigation.

Acknowledgments

This study was funded by institutional startup funds (A.S. and P.M.W.). This work was supported in part by National Institutes of Health; National Cancer Institute Grant R01CA136665 (to J.A.C. and R.C.S.); Florida Department of Health Bankhead-Coley Cancer Research Program [Grants FL09B202 (to J.A.C. and R.C.S.) and FL3BF01 (to J.A.C.)].

Author Disclosure Statement

No competing financial interests exist.

References

1. Keutgen XM, Sadowski SM, Kebebew E 2015 Management of anaplastic thyroid cancer. Gland Surg **4:**44–51.
2. Smallridge RC, Copland JA, Brose MS, Wadsworth JT, Houvras Y, Menefee ME, Bible KC, Shah MH, Gramza AW, Klopper JP, Marlow LA, Heckman MG, Von Roemeling R 2013 Efatutazone, an oral PPAR-gamma agonist, in combination with paclitaxel in anaplastic thyroid cancer: results of a multicenter phase 1 trial. J Clin Endocrinol Metab **98:**2392–2400.
3. Schoenfeld JD, Odejide OO, Wirth LJ, Chan AW 2012 Survival of a patient with anaplastic thyroid cancer following intensity-modulated radiotherapy and sunitinib: a case report. Anticancer Res **32:**1743–1746.
4. McLarnon A 2012 Thyroid cancer: Pazopanib alone is not effective against anaplastic thyroid cancer. Nat Rev Endocrinol **8:**565.
5. Antonelli A, Fallahi P, Ulisse S, Ferrari SM, Minuto M, Saraceno G, Santini F, Mazzi V, D'Armiento M, Miccoli P 2012 New targeted therapies for anaplastic thyroid cancer. Anticancer Agents Med Chem **12:**87–93.
6. Brown RF, Ducic Y 2013 Aggressive surgical resection of anaplastic thyroid carcinoma may provide long-term survival in selected patients. Otolaryngol Head Neck Surg **148:**564–571.
7. Foote RL, Molina JR, Kasperbauer JL, Lloyd RV, McIver B, Morris JC, Grant CS, Thompson GB, Richards ML, Hay ID, Smallridge RC, Bible KC 2011 Enhanced survival in locoregionally confined anaplastic thyroid carcinoma: a single-institution experience using aggressive multimodal therapy. Thyroid **21:**25–30.
8. Sharma A, Podolsky R, Zhao J, McIndoe RA 2009 A modified hyperplane clustering algorithm allows for efficient and accurate clustering of extremely large datasets. Bioinformatics **25:**1152–1157.
9. Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA 2003 DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol **4:**P3.
10. Kramer A, Green J, Pollard J, Jr., Tugendreich S 2014 Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics **30:**523–530.
11. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS 2009 MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res **37:**W202–208.
12. Marlow LA, D'Innocenzi J, Zhang Y, Rohl SD, Cooper SJ, Sebo T, Grant C, McIver B, Kasperbauer JL, Wadsworth JT, Casler JD, Kennedy PW, Highsmith WE, Clark O, Milosevic D, Netzel B, Cradic K, Arora S, Beaudry C, Grebe SK, Silverberg ML, Azorsa DO, Smallridge RC, Copland JA 2010 Detailed molecular fingerprinting of four new anaplastic thyroid carcinoma cell lines and their use for verification of RhoB as a molecular therapeutic target. J Clin Endocrinol Metab **95:**5338–5347.
13. Qiu J, Zhang W, Xia Q, Liu F, Li L, Zhao S, Gao X, Zang C, Ge R, Sun Y 2015 RNA sequencing identifies crucial genes in papillary thyroid carcinoma (PTC) progression. Exp Mol Pathol **100:**151–159.

14. Palona I, Namba H, Mitsutake N, Starenki D, Podtcheko A, Sedliarou I, Ohtsuru A, Saenko V, Nagayama Y, Umezawa K, Yamashita S 2006 BRAFV600E promotes invasiveness of thyroid cancer cells through nuclear factor kappaB activation. Endocrinology **147:**5699–5707.
15. Rodrigues RF, Roque L, Krug T, Leite V 2007 Poorly differentiated and anaplastic thyroid carcinomas: chromosomal and oligo-array profile of five new cell lines. Br J Cancer **96:**1237–1245.
16. Hebrant A, Dom G, Dewaele M, Andry G, Tresallet C, Leteurtre E, Dumont JE, Maenhaut C 2012 mRNA expression in papillary and anaplastic thyroid carcinoma: molecular anatomy of a killing switch. PLoS One **7:**e37807.
17. Smallridge RC, Marlow LA, Copland JA 2009 Anaplastic thyroid cancer: molecular pathogenesis and emerging therapies. Endocr Relat Cancer **16:**17–44.
18. Fiumara A, Belfiore A, Russo G, Salomone E, Santonocito GM, Ippolito O, Vigneri R, Gangemi P 1997 In situ evidence of neoplastic cell phagocytosis by macrophages in papillary thyroid cancer. J Clin Endocrinol Metab **82:**1615–1620.
19. Ryder M, Ghossein RA, Ricarte-Filho JC, Knauf JA, Fagin JA 2008 Increased density of tumor-associated macrophages is associated with decreased survival in advanced thyroid cancer. Endocr Relat Cancer **15:**1069–1074.
20. Caillou B, Talbot M, Weyemi U, Pioche-Durieu C, Al Ghuzlan A, Bidart JM, Chouaib S, Schlumberger M, Dupuy C 2011 Tumor-associated macrophages (TAMs) form an interconnected cellular supportive network in anaplastic thyroid carcinoma. PLoS One **6:**e22567.
21. von Roemeling CA, Marlow LA, Pinkerton AB, Crist A, Miller J, Tun HW, Smallridge RC, Copland JA 2015 Aberrant lipid metabolism in anaplastic thyroid carcinoma reveals stearoyl CoA desaturase 1 as a novel therapeutic target. J Clin Endocrinol Metab **100:**E697–709.
22. Rodriguez-Rodero S, Fernandez AF, Fernandez-Morera JL, Castro-Santos P, Bayon GF, Ferrero C, Urdinguio RG, Gonzalez-Marquez R, Suarez C, Fernandez-Vega I, Fresno Forcelledo MF, Martinez-Camblor P, Mancikova V, Castelblanco E, Perez M, Marron PI, Mendiola M, Hardisson D, Santisteban P, Riesco-Eizaguirre G, Matias-Guiu X, Carnero A, Robledo M, Delgado-Alvarez E, Menendez-Torre E, Fraga MF 2013 DNA methylation signatures identify biologically distinct thyroid cancer subtypes. J Clin Endocrinol Metab **98:**2811–2821.
23. Onda M, Emi M, Yoshida A, Miyamoto S, Akaishi J, Asaka S, Mizutani K, Shimizu K, Nagahama M, Ito K, Tanaka T, Tsunoda T 2004 Comprehensive gene expression profiling of anaplastic thyroid cancers with cDNA microarray of 25 344 genes. Endocr Relat Cancer **11:**843–854.
24. Lee JJ, Au AY, Foukakis T, Barbaro M, Kiss N, Clifton-Bligh R, Staaf J, Borg A, Delbridge L, Robinson BG, Wallin G, Hoog A, Larsson C 2008 Array-CGH identifies cyclin D1 and UBCH10 amplicons in anaplastic thyroid carcinoma. Endocr Relat Cancer **15:**801–815.
25. Jeon MJ, Chun SM, Kim D, Kwon H, Jang EK, Kim TY, Kim WB, Shong YK, Jang SJ, Song DE, Kim WG 2016 Genomic alterations of anaplastic thyroid carcinoma detected by targeted massive parallel sequencing in a BRAF(V600E) mutation-prevalent area. Thyroid **26:**683–690.
26. Landa I, Ibrahimpasic T, Boucai L, Sinha R, Knauf JA, Shah RH, Dogan S, Ricarte-Filho JC, Krishnamoorthy GP, Xu B, Schultz N, Berger MF, Sander C, Taylor BS, Ghossein R, Ganly I, Fagin JA 2016 Genomic and transcriptomic hallmarks of poorly differentiated and anaplastic thyroid cancers. J Clin Invest **126:**1052–1066.
27. Sykorova V, Dvorakova S, Vcelak J, Vaclavikova E, Halkova T, Kodetova D, Lastuvka P, Betka J, Vlcek P, Reboun M, Katra R, Bendlova B 2015 Search for new genetic biomarkers in poorly differentiated and anaplastic thyroid carcinomas using next generation sequencing. Anticancer Res **35:**2029–2036.
28. Kunstman JW, Juhlin CC, Goh G, Brown TC, Stenman A, Healy JM, Rubinstein JC, Choi M, Kiss N, Nelson-Williams C, Mane S, Rimm DL, Prasad ML, Hoog A, Zedenius J, Larsson C, Korah R, Lifton RP, Carling T 2015 Characterization of the mutational landscape of anaplastic thyroid cancer via whole-exome sequencing. Hum Mol Genet **24:**2318–2329.
29. Hebrant A, Floor S, Saiselet M, Antoniou A, Desbuleux A, Snyers B, La C, de Saint Aubain N, Leteurtre E, Andry G, Maenhaut C 2014 miRNA expression in anaplastic thyroid carcinomas. PLoS One **9:**e103871.
30. Chen C, Grennan K, Badner J, Zhang D, Gershon E, Jin L, Liu C 2011 Removing batch effects in analysis of expression microarray data: an evaluation of six batch adjustment methods. PLoS One **6:**e17238.
31. Nygaard V, Rodland EA, Hovig E 2016 Methods that remove batch effects while retaining group differences may lead to exaggerated confidence in downstream analyses. Biostatistics **17:**29–39.
32. Buehler D, Hardin H, Shan W, Montemayor-Garcia C, Rush PS, Asioli S, Chen H, Lloyd RV 2013 Expression of epithelial-mesenchymal transition regulators SNAI2 and TWIST1 in thyroid carcinomas. Mod Pathol **26:**54–61.
33. Montero-Conde C, Martin-Campos JM, Lerma E, Gimenez G, Martinez-Guitarte JL, Combalia N, Montaner D, Matias-Guiu X, Dopazo J, de Leiva A, Robledo M, Mauricio D 2008 Molecular profiling related to poor prognosis in thyroid carcinoma. Combining gene expression data and biological information. Oncogene **27:**1554–1561.
34. Espinal-Enriquez J, Munoz-Montero S, Imaz-Rosshandler I, Huerta-Verde A, Mejia C, Hernandez-Lemus E 2015 Genome-wide expression analysis suggests a crucial role of dysregulation of matrix metalloproteinases pathway in undifferentiated thyroid carcinoma. BMC Genomics **16:**207.
35. Knauf JA, Sartor MA, Medvedovic M, Lundsmith E, Ryder M, Salzano M, Nikiforov YE, Giordano TJ, Ghossein RA, Fagin JA 2011 Progression of BRAF-induced thyroid cancer is associated with epithelial-mesenchymal transition requiring concomitant MAP kinase and TGFbeta signaling. Oncogene **30:**3153–3162.
36. Hsu YL, Hou MF, Kuo PL, Huang YF, Tsai EM 2013 Breast tumor-associated osteoblast-derived CXCL5 increases cancer progression by ERK/MSK1/Elk-1/snail signaling pathway. Oncogene **32:**4436–4447.
37. Lim JB, Chung HW 2015 Serum ENA78/CXCL5, SDF-1/CXCL12, and their combinations as potential biomarkers for prediction of the presence and distant metastasis of primary gastric cancer. Cytokine **73:**16–22.
38. Liao W, Liu W, Yuan Q, Liu X, Ou Y, He S, Yuan S, Qin L, Chen Q, Nong K, Mei M, Huang J 2013 Silencing of DLGAP5 by siRNA significantly inhibits the proliferation and invasion of hepatocellular carcinoma cells. PLoS One **8:**e80789.
39. Sabatier R, Finetti P, Adelaide J, Guille A, Borg JP, Chaffanet M, Lane L, Birnbaum D, Bertucci F 2011 Down-

regulation of ECRG4, a candidate tumor suppressor gene, in human breast cancer. PLoS One **6:**e27656.
40. Li LW, Yu XY, Yang Y, Zhang CP, Guo LP, Lu SH 2009 Expression of esophageal cancer related gene 4 (ECRG4), a novel tumor suppressor gene, in esophageal cancer and its inhibitory effect on the tumor growth in vitro and in vivo. Int J Cancer **125:**1505–1513.
41. Vizcaino C, Mansilla S, Portugal J 2015 Sp1 transcription factor: A long-standing target in cancer chemotherapy. Pharmacol Ther **152:**111–124.
42. Qi M, Li Y, Liu J, Yang X, Wang L, Zhou Z, Han B 2012 Morphologic features of carcinomas with recurrent gene fusions. Advances in anatomic pathology **19:**417–424.
43. Teixeira MR 2006 Recurrent fusion oncogenes in carcinomas. Crit Rev Oncog **12:**257–271.
44. Prasad ML, Vyas M, Horne MJ, Virk RK, Morotti R, Liu Z, Tallini G, Nikiforova MN, Christison-Lagay ER, Udelsman R, Dinauer CA, Nikiforov YE 2016 NTRK fusion oncogenes in pediatric papillary thyroid carcinoma in northeast United States. Cancer **122:**1097–1107.
45. Wang J, Li L, Liu S, Zhao Y, Wang L, Du G 2016 FOXC1 promotes melanoma by activating MST1R/PI3K/AKT. Oncotarget. [Epub ahead of print]; DOI: 10.18632/oncotarget .11224.
46. Yong KJ, Gao C, Lim JS, Yan B, Yang H, Dimitrov T, Kawasaki A, Ong CW, Wong KF, Lee S, Ravikumar S, Srivastava S, Tian X, Poon RT, Fan ST, Luk JM, Dan YY, Salto-Tellez M, Chai L, Tenen DG 2013 Oncofetal gene SALL4 in aggressive hepatocellular carcinoma. N Engl J Med **368:**2266–2276.
47. Kruiswijk F, Hasenfuss SC, Sivapatham R, Baar MP, Putavet D, Naipal KA, van den Broek NJ, Kruit W, van der Spek PJ, van Gent DC, Brenkman AB, Campisi J, Burgering BM, Hoeijmakers JH, de Keizer PL 2015 Targeted inhibition of metastatic melanoma through interference with Pin1-FOXM1 signaling. Oncogene 35:2166–2177.
48. Pobbati AV, Han X, Hung AW, Weiguang S, Huda N, Chen GY, Kang C, Chia CS, Luo X, Hong W, Poulsen A 2015 Targeting the central pocket in human transcription factor TEAD as a potential cancer therapeutic strategy. Structure **23:**2076–2086.
49. Elumalai N, Berg A, Rubner S, Berg T 2015 Phosphorylation of capsaicinoid derivatives provides highly potent and selective inhibitors of the transcription factor STAT5b. ACS Chem Biol **10:**2884–2890.
50. Pyonteck SM, Akkari L, Schuhmacher AJ, Bowman RL, Sevenich L, Quail DF, Olson OC, Quick ML, Huse JT, Teijeiro V, Setty M, Leslie CS, Oei Y, Pedraza A, Zhang J, Brennan CW, Sutton JC, Holland EC, Daniel D, Joyce JA 2013 CSF-1R inhibition alters macrophage polarization and blocks glioma progression. Nat Med **19:**1264–1272.

Address correspondence to:
Ashok Sharma, PhD
Center for Biotechnology and Genomic Medicine
Medical College of Georgia at Augusta University
1120 15th Street, CA-4094
Augusta, GA 30912

E-mail: assharma@augusta.edu