Transposon mutagenesis identifies chromatin modifiers cooperating with *Ras* in thyroid tumorigenesis and detects *ATXN7* as a cancer gene

Cristina Montero-Conde[a,b,1], Luis J. Leandro-Garcia[a,1], Xu Chen[a], Gisele Oler[a], Sergio Ruiz-Llorente[a], Mabel Ryder[a], Iñigo Landa[a], Francisco Sanchez-Vega[c], Konnor La[c], Ronald A. Ghossein[d], Dean F. Bajorin[e], Jeffrey A. Knauf[a], Jesse D. Riordan[f], Adam J. Dupuy[f], and James A. Fagin[a,e,2]

[a]Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065; [b]Hereditary Endocrine Cancer Group, Spanish National Cancer Research Center, Madrid, Spain 28029; [c]Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center, New York, NY 10065; [d]Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY 10065; [e]Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY 10065; and [f]Department of Anatomy and Cell Biology, University of Iowa, Iowa City, IA 52242

Edited by Albert de la Chapelle, Ohio State University Comprehensive Cancer Center, Columbus, OH, and approved May 16, 2017 (received for review February 16, 2017)

Oncogenic *RAS* mutations are present in 15–30% of thyroid carcinomas. Endogenous expression of mutant Ras is insufficient to initiate thyroid tumorigenesis in murine models, indicating that additional genetic alterations are required. We used Sleeping Beauty (SB) transposon mutagenesis to identify events that cooperate with $Hras^{G12V}$ in thyroid tumor development. Random genomic integration of SB transposons primarily generated loss-of-function events that significantly increased thyroid tumor penetrance in *Tpo-Cre/homozygous FR-$Hras^{G12V}$* mice. The thyroid tumors closely phenocopied the histological features of human RAS-driven, poorly differentiated thyroid cancers. Characterization of transposon insertion sites in the SB-induced tumors identified 45 recurrently mutated candidate cancer genes. These mutation profiles were remarkably concordant with mutated cancer genes identified in a large series of human poorly differentiated and anaplastic thyroid cancers screened by next-generation sequencing using the MSK-IMPACT panel of cancer genes, which we modified to include all SB candidates. The disrupted genes primarily clustered in chromatin remodeling functional nodes and in the PI3K pathway. *ATXN7*, a component of a multiprotein complex with histone acetylase activity, scored as a significant SB hit. It was recurrently mutated in advanced human cancers and significantly co-occurred with *RAS* or *NF1* mutations. Expression of *ATXN7* mutants cooperated with oncogenic RAS to induce thyroid cell proliferation, pointing to *ATXN7* as a previously unrecognized cancer gene.

thyroid cancer genomics | Sleeping Beauty | Ras | Swi/Snf | Pten

Oncogenic *RAS* mutations are believed to arise early in the course of thyroid tumor development (1). They are present in 13% of well-differentiated thyroid cancers (2) and are enriched in advanced forms of the disease (3). However, physiological expression of oncogenic *Ras* alleles is not sufficient to trigger thyroid transformation in mice (4–8), indicating that other cooperative molecular events are required to promote tumor development.

The Thyroid Cancer (THCA) Network of The Cancer Genome Atlas (TCGA) reported a comprehensive integrated analysis of the genomic landscape of papillary thyroid carcinomas, a low-grade tumor with a low mutation burden and infrequent copy number abnormalities. This analysis identified few genomic interactions with *RAS* other than a higher frequency of chromosome 22q loss (2). Recent genomic studies of poorly differentiated thyroid cancers (PDTCs) and anaplastic thyroid cancers (ATCs), which are virulent forms of the disease associated with a higher frequency of genetic alterations, again revealed the co-occurrence of *RAS* mutations with 22q loss of heterozygosity, as well as mutations of *TP53*, the *TERT* promoter, *EIF1AX*, and *PTEN* (3, 9). Less frequently, PDTCs and ATCs are associated with mutations of *NF2*, as well as of genes encoding components of the SWI/SNF chromatin remodeling complex, histone methyltransferases, and other genes involved in epigenetic processes (3). Other than for *NF2* (10), there is limited information about the possible functional relevance of these RAS-associated genetic events in the context of thyroid cancer.

The Sleeping Beauty (SB) transposon-based mutagenesis system is an unbiased approach to capture functional events that promote cancer in the mouse. It has been used to identify novel cancer genes in several tumor types (11–14) and to provide insights into combinations of gene alterations that induce transformation in particular cellular contexts (15, 16). Here we used the SB system to introduce insertional disruptions randomly into the genome of $Hras^{G12V}$ mutant thyroid follicular cells to discover genes that cooperate to drive tumor development, because $Hras^{G12V}$ alone is insufficient to promote transformation. We identified 45 candidate cancer genes using a genecentric common insertion site (gCIS), a statistical method that identifies National Center of Biotechnical Information Reference Sequence (RefSeq) genes in

Significance

Mutations of RAS are believed to be early events in thyroid tumorigenesis but are insufficient to induce transformation. A forward genetic screen with transposons engineered to integrate randomly into the mouse Ras-mutant thyroid cell genome and to disrupt genes at their insertion sites resulted in tumors that phenocopied human RAS-driven, poorly differentiated thyroid cancers. Analysis of the transposon-integration sites revealed recurrent mutations of chromatin modifiers and PI3K pathway genes, consistent with mutations seen in human advanced thyroid cancers. These human cancers have a high mutation burden, which confounds distinctions between driver and passenger mutations. This unbiased screen for genes selected during tumorigenesis provides strong functional support for genetic disruptions in these pathways in RAS-induced thyroid tumor progression.

Author contributions: C.M.-C., L.J.L.-G., X.C., S.R.-L., I.L., J.A.K., J.D.R., A.J.D., and J.A.F. designed research; C.M.-C., L.J.L.-G., X.C., G.O., S.R.-L., M.R., I.L., R.A.G., J.D.R., and A.J.D. performed research; J.D.R. and A.J.D. contributed new reagents/analytic tools; C.M.-C., L.J.L.-G., F.S.-V., K.L., R.A.G., D.F.B., J.A.K., J.D.R., A.J.D., and J.A.F. analyzed data; and C.M.-C., L.J.L.-G., and J.A.F. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The bulk of MSK-IMPACT sequencing data (Landa I et al., ref. 3) is publicly available at the Memorial Sloan Kettering BioPortal for Cancer Genomics (www.cbioportal.org/study?id=thyroid_mskcc_2016#summary). All other sequencing data are included in Datasets S1 and S4.

[1]C.M.-C. and L.J.L.-G. contributed equally to this study.

[2]To whom correspondence should be addressed. Email: faginj@mskcc.org.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1702723114/-/DCSupplemental.

the tumor series with a higher frequency of transposon-mediated disruption than predicted based on a random pattern of integration in the absence of selection (17). These genes primarily clustered into two molecular networks: chromatin remodeling and serine-threonine protein kinases, which are predominantly effectors in the PI3K pathway. Deep sequencing of SB candidate genes in a panel of 117 advanced human thyroid cancers found that 30 of 45 SB hits were mutated, with *ARID1B*, *ARID2,* and *ATXN7* being the most frequently disrupted genes. The strong concordance of hits in this forward genetic screen with genetic events and specific functional nodes in human thyroid cancers supports their role in thyroid tumorigenesis. As further validation we explored the functional consequence of *ATXN7* mutations, a component of the SPT3/TAF9/GCN5 histone acetyltransferase (SAGA) complex, because this SB hit has not been previously implicated in cancer. The somatic *ATXN7* thyroid cancer mutants clustered in the polyglutamine domain of the protein, the germline expansion of which causes the autosomal dominant syndrome spinocerebellar ataxia 7. Expression of ATXN7 mutants induced proliferation in the context of oncogenic RAS in thyroid cells, pointing to a role of this ubiquitously expressed protein in cancer.

Results

The SB System Induces Thyroid Tumorigenesis in the Context of Hras^{G12V}. We used the SB system to identify genetic events that cooperate with Hras^{G12V} to drive thyroid transformation in mice. This model is based on the ability of engineered transposons to integrate randomly between or into genes and to introduce gain- or loss-of-function alterations (18). The approach requires the expression of a transposase, which catalyzes transposon mobilization into largely random sites in the mouse genome. We crossed the Rosa LSL-SB transposase/T6113 transgenic mouse line, which harbors the transposon concatemer on chromosome 1 (19), with *Tpo-Cre/FR-Hras^{G12V}* mice to direct the expression of the SB transposase to thyroid follicular cells in the context of mutant *Hras*. Only 2% (2/90) of the *Tpo-Cre/FR-Hras^{Hom}* mice (5), and none of the triple-transgenic *Tpo-Cre/FR-Hras^{Het}/Rosa LSL-SB* or *Tpo-Cre/Rosa LSL-SB/T6113* mice, developed thyroid nodules by 17 mo of age. By contrast, 16% (22/139) of the quadruple-transgenic *Tpo-Cre/FR-Hras^{Hom}/Rosa LSL-SB/T6113* mice developed at least one nodule in the thyroid gland by age 15 mo (unpaired *t* test, *P* value <0.001 vs. *Tpo-Cre/FR-Hras^{Hom}* mice) (Table 1). Interestingly, there was no transposon-driven increase in tumor frequency in the context of *Hras^{G12V}* heterozygosity, a finding that is consistent with the requirement for copy number imbalances of the oncogenic *Ras* allele to induce transformation (5, 20, 21).

Because of their small size, most tumor tissues were used entirely for DNA extraction, but there was sufficient material for histological evaluation in 6 of 26 tumors from the quadruple-transgenic mice and from the two nodules that developed in the aged *Tpo-Cre/Hras^{Hom}* controls. Five of six nodules from the quadruple transgenics had the typical histological characteristics of encapsulated PDTC (Fig. 1 *A–C*), i.e., cells growing in a solid growth pattern, frequent mitotic figures, and the presence of necrosis. There also was capsular invasion, all features that phenocopy human PDTCs driven by oncogenic *RAS* mutations (3, 22, 23). The remaining tumor was an encapsulated, well-differentiated papillary carcinoma. The two tumors arising in the *Tpo-Cre/Hras^{Hom}* controls were well-differentiated thyroid neoplasms and were likely benign.

SB Insertion-Site Analysis Reveals Putative Genes and Networks That Cooperate with Hras^{G12V} in Transformation. Nodules were either macroscopically dissected or laser-capture microdissected from the mouse thyroid to isolate tumor DNA, which then was digested by specific restriction enzymes followed by a ligation-mediated PCR with barcoded oligos against the transposon–genomic DNA (gDNA) junction sites (18). We then performed high-throughput sequencing to localize clonally expanded insertion sites (see *Materials and Methods* for details) in the genome of the mouse for each nodule. Of the 26 nodules, 7 were not informative: The number of mapped reads was too low to call clonal insertion sites in three nodules, and four others did not show any clonal insertion site into a RefSeq gene (Dataset S1). We used the gCIS analysis for the remaining 19 tumors to identify genomic regions within or near annotated genes harboring transposon insertions at a frequency higher than expected by chance, indicative of positive selection and clonal expansion of the affected cells within the growing tumor mass (17). Only genes with transposon insertions in at least three different tumors were considered. Based on this threshold, we identified 45 gCISs (Fig. 1*D*). Seventeen of the 19 tumors harbored at least one gCIS, with a median of 10 gCIS per tumor (range, 1–31). *Pten* was the most frequently targeted SB hit (10/17 nodules), followed by *Smg1*, *Kmt2e*, and *Magi1* (Fig. 1*D*). Functional annotation clustering of SB genes using the DAVID database (24, 25) identified the terms chromatin modification (P = 5.4E-6, Benjamini-corrected,1.7E-3), SET protein domain (P = 2.7E-4, Benjamini-corrected, 3.7E-2), and serine/threonine protein kinase (P = 2.1E-3, Benjamini-corrected, 4.6E-02) among the most significantly enriched terms when using a high classification stringency (similarity threshold = 0.85) and a Fisher exact P value ≤0.01 for term enrichment (Fig. 1*E* and Dataset S2). In addition, STRING functional protein association analysis revealed two protein network nodes, one related to chromatin remodeling and the second to the PI3K signaling pathway (medium confidence, 0.4–0.7) (Fig. 1*F*).

Frequent Mutation of the PI3K Signaling Pathway in Hras^{G12V} Thyroid Follicular Cells. The canonical signaling cascades of *Hras^{G12V}* include Raf/Mek/Erk and Pi3k/Akt. To determine if these signaling axes were activated in homozygous *Hras^{G12V}* thyroid follicular cells, we performed immunohistochemical (IHC) staining for phospho-ERK and phospho-AKT in representative thyroid sections of *Tpo-Cre/FR-Hras^{Hom}* mice (Fig. 2*B*). Both effectors were modestly activated in comparison with *Hras^{wt}* thyroid follicular cells (in *Tpo-Cre* mice) but at levels that were insufficient to trigger

Table 1. The SB system drives thyroid tumorigenesis in the context of Hras^{G12V}

Genotype	Total no. mice	No. mice/no. nodules	Mice, %	Age in wk (range)
Tpo-Cre/FR-Hras^{Het}	25	0	0	72.4 (60–139)
Tpo-Cre/FR-Hras^{Hom}	90	2/2	2.2*	74.5 (71–78)
Tpo-Cre/FR-Hras^{Het}/Rosa LsL-SB	25	0	0	64.5 (52–75)
Tpo-Cre/Rosa LsL-SB/T6113	8	0	0	69.5 (62–75)
Tpo-Cre/FR-Hras^{Het}/Rosa LsL-SB/T6113	58	0	0	64.8 (59–81)
Tpo-Cre/FR-Hras^{Hom}/Rosa LsL-SB/T6113	139	22/26	15.8*	66.75 (56–89)

Number and percentage of transgenic mice with thyroid nodules and age at death for each genotype cohort are shown.
*P < 0.001.

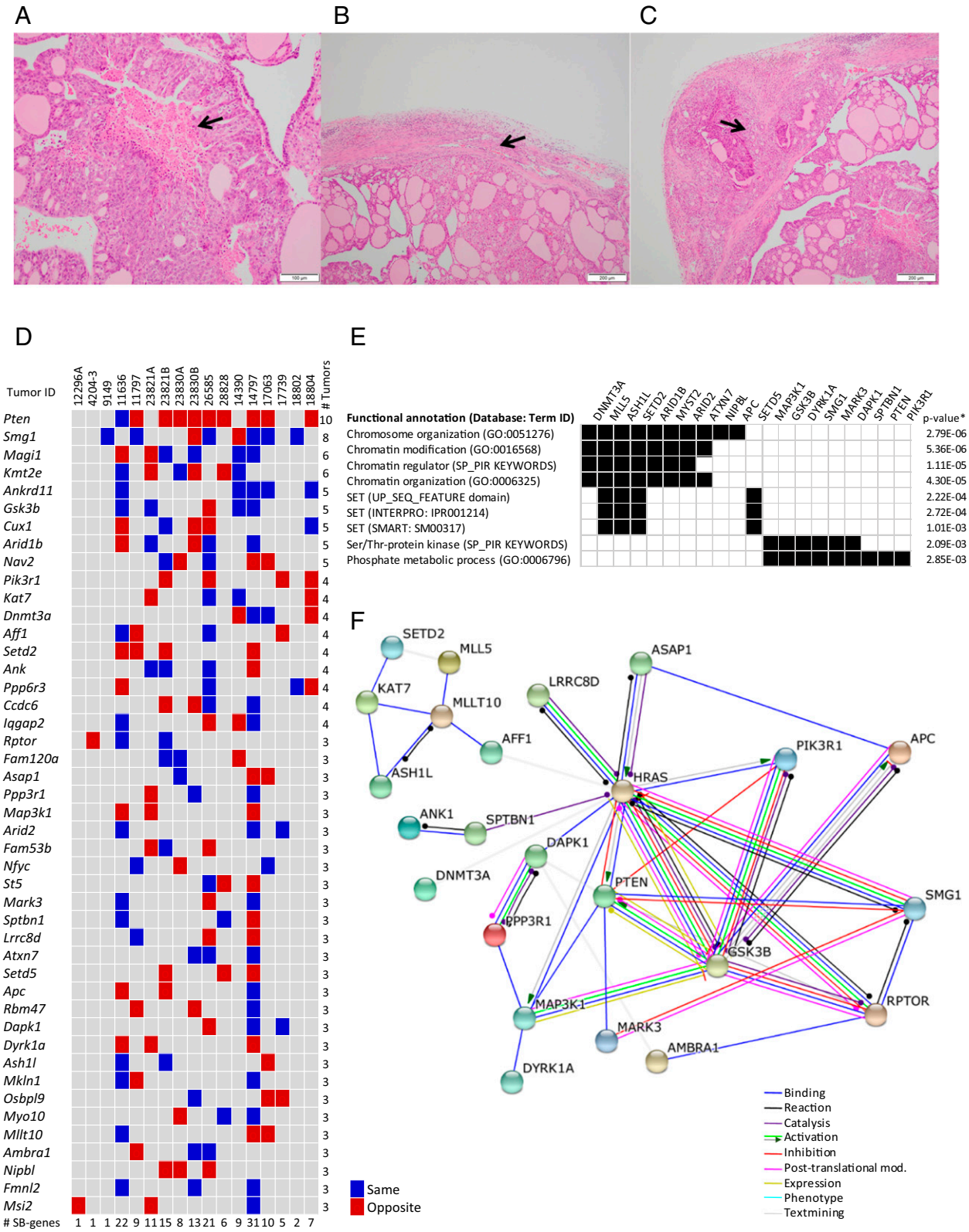

Fig. 1. The SB system induces tumors that phenocopy human PTDCs and identifies genes and functional nodes that cooperate with $Hras^{G12V}$ in thyroid tumorigenesis. (*A–C*) Representative photomicrographs of H&E-stained sections of thyroid tumors. Arrows point to distinct features consistent with PDTC: necrosis (*A*), encapsulation (*B*), and capsular invasion by tumoral cells (*C*). (Scale bars: 100 μm in *A*, 200 μm in *B* and C.) (*D*) Oncoprint of SB candidate genes identified by the gCIS analysis of SB-induced tumors. Transposon insertions in the same sense as gene transcription are highlighted in blue, and those in the opposite sense are in red. (*E*) Functional annotation analysis using the DAVID tool suite. *The asterisk indicates the modified Fisher exact *P* value (DAVID EASE score) for SB gene-enrichment in the indicated annotation terms. GO: Gene Ontology database; SP_PIR keywords: SwissProt/Uniprot and PIR (Protein Information Resource) databases; UP_SEQ feature domain: sites of interest in the protein sequence from the UniProt database. (*F*) Interaction networks of SB hits according to gene-coding protein–protein interactions, gene coexpression, phenotype, and text mining (co-occurrence in abstracts) using the STRING functional protein association database.

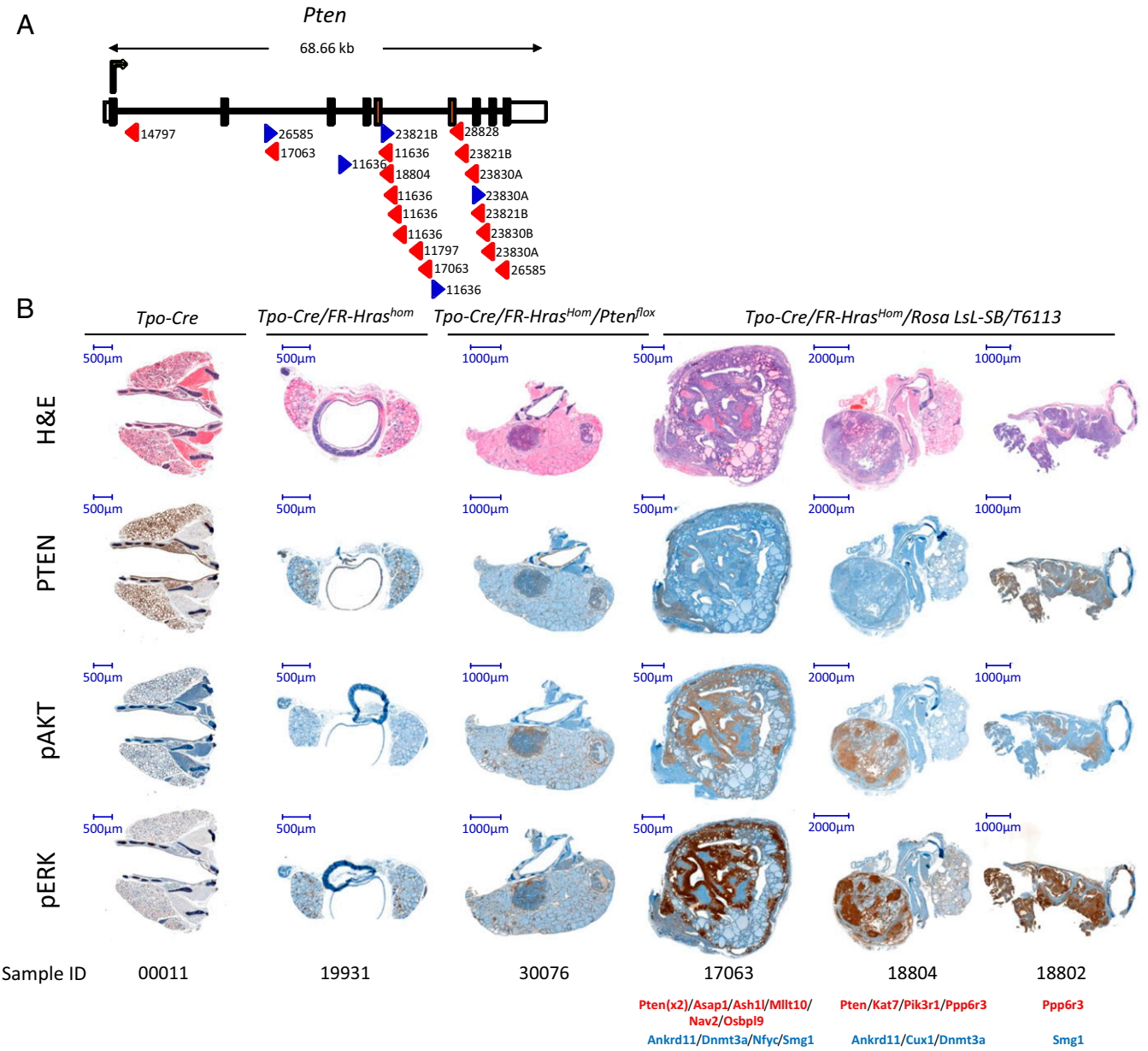

Fig. 2. *Pten* is a major SB hit in the context of $Hras^{G12V}$ and results in thyroid tumors with activation of PI3K–AKT signaling. (*A*) Map of transposon clonal integration sites into the *Pten* gene sequence in the SB-induced thyroid tumors. Blue and red arrowheads indicate transposon insertions integrated in the sense or antisense direction of gene transcription, respectively. (*B*) Serial sections of representative thyroid tissues from WT mice (Tpo-Cre, *Tpo-Cre/FR-*$Hras^{Hom}$, *Tpo-Cre/FR-*$Hras^{Hom}$*/*$Pten^{flox}$) and of three tumors from *Tpo-Cre/FR-*$Hras^{Hom}$*/Rosa LSL-SB/T6113* mice stained with H&E or by IHC for Pten, phospho-Akt^{Ser473}, and phospho-ERK. Transposon gCISs corresponding to each of the SB-induced tumors are shown below. Red indicates antisense integration; blue indicates sense integration.

transformation. A recent study showed that loss of either *Nf2* or *Pten* in $Hras^{Hom}$ thyroid follicular cells, associated with augmentation of Mapk or Pi3K-Akt signaling, respectively, promoted tumor development, (10). Remarkably, 76.5% (13/17) of SB-induced tumors with a gCIS harbored at least one clonal insertion site in negative regulators of the PI3K pathway: *Pten* (10/17, 58.8%), *Pik3r1* (4/17, 23.5%), and *Magi1* (6/17, 35.3%). Transposon insertions were primarily oriented inversely (*Pten* and *Pik3r1*) or randomly (*Magi1*) to the direction of transcription (Figs. 1*D* and 2*A*), suggesting loss of function. IHC staining of Pten and phospho-Akt showed loss of Pten expression with reciprocal increase in phospho-Akt in tumors with gCIS within the *Pten* gene. Other PI3K-related genes frequently targeted by SB transposition included the mTOR regulatory-associated protein-encoding gene *Rptor* (3/17, 17.6%), the PI3K-related kinase gene *Smg1* (8/17, 47%), and the PI3K scaffolding protein-encoding gene *Fam120a* (3/17, 17.6%) (26). Interestingly, IHC for pERK was dramatically augmented compared with *Tpo-Cre/FR-*$Hras^{Hom}$ controls, pointing to the likely acquisition of other events besides *Pten* loss during transformation (Fig. 2*B*).

In view of the high frequency of tumor-associated clonal *SB* integration events into genes encoding effectors of the PI3K pathway, and particularly within the *Pten* gene, we recently investigated this interaction by crossing *Tpo-Cre/*$Hras^{Hom}$ or *Tpo-Cre/*$Hras^{Het}$ mice with $Pten^{flox/flox}$ mice. Compound mice developed encapsulated tumors that phenocopied PDTC and that led to decreased survival, primarily in the context of $Hras^{G12V}$ homozygosity, confirming the significance of this association (10).

We next validated the expression of predicted chimeric transposon/Ankrd11 gene transcripts because, unlike *Pten*, all transposons within this gene integrated in the direction of transcription (Fig. S1). The *Ankrd11* gene is believed to function within the cohesin complex (27), other components of which (e.g., *Med12*) are known to be mutated in advanced human thyroid cancers. We performed RT-PCR with primers bracketing a site within the transposon with sites distal to each distinct integration site on the gene in four tumors and successfully amplified appropriate products in two of them, which were confirmed by Sanger sequencing (Fig. S1). Interestingly, analysis of *ANKRD11* sequencing defects in the pan-cancer Memorial Sloan Kettering (MSK) cBioPortal database identified a subset of tumors harboring intragenic deletions or duplications upstream of exon 9, some of which may point to unrecognized fusions (this interpretation is unconfirmed, because intron 8 of *ANKRD11* was not sequenced in the MSK-IMPACT panel) (Fig. S1).

SB Genes Are Frequently Mutated in Human Thyroid Cancer. The set of 45 SB candidates is significantly enriched in genes whose human homologs are in the Cancer Gene Census maintained by the Sanger Institute (cancer.sanger.ac.uk/census) (28). Altogether, 13 of 45 (29%) SB genes (*AFF1*, *APC*, *ARID1B*, *ARID2*, *CCDC6*, *CUX1*, *DNMT3A*, *MAP3K1*, *MLLT10*, *MSI2*, *PIK3R1*, *PTEN*, and *SETD2*) are included in this census, compared with 616 of 19,066 (3%) protein-coding genes in the human genome (P = 1e-9, one-sided Fisher's exact test).

Moreover, the human homologs of 25 of the 45 SB candidate genes (55%) had at least one missense or truncating mutation detected by whole-exome sequencing in the papillary thyroid cancer THCA-TCGA series (n = 402) (Fig. 3*A*) (2). This enrichment is highly significant based on a random permutation experiment in which sets of 45 genes were selected at random from all human protein-coding genes (P < 1E-6; see *Materials and Methods* for details). The most recurrently mutated gCIS homolog was *ARID1B* (5/402 tumors, 1.24%) followed by *SETD5*, *DNMT3A*, and *ASHL1* (all 3/402, 0.75%). *PTEN* mutations were seen in 2 of 402 tumors.

The THCA-TCGA study focused only on papillary thyroid carcinomas and excluded advanced forms of the disease [PDTC and ATC] (2). To explore whether SB-targeted genes were recurrently mutated in these cancers, we mined the data of a recently reported next-generation sequencing profile of 117 PDTCs and ATCs. This study used the MSK-IMPACT cancer exome panel (3), a platform that captures all exons and selected introns of 341 cancer genes (29). Because only 11 of 45 SB hits within protein-coding genes are included in MSK-IMPACT, we designed a custom Nimblegen sequencing panel to identify somatic mutations and copy number alterations in the remaining 34 SB genes. Eighty-two of the 117 advanced carcinomas (71%) were analyzed with the version of the MSK-IMPACT augmented to include all SB candidate genes (Fig. 3*B*). Sixty-seven percent (30/45) of SB genes showed at least one missense or truncating mutation in this tumor series, constituting a statistically significant level of enrichment based on a random permutation experiment (P = 0.0046).

The most recurrently mutated genes were *ATXN7* (6/82, 7.3%), *PTEN* (8/117, 6.8%), *ARID1B* (8/117, 6.8%), and *ARID2* (5/117, 4.3%) followed by *MAGI1* and *KMT2E* (both 3/82, 3.7%) (Dataset S4). The frequency of RAS mutations was 33 of 117 (28.2%) (*NRAS* = 25, *HRAS* = 6, and *KRAS* = 2) in the entire series and 21 of 82 (25.6%) in the 82-sample subset (*NRAS* = 15, *HRAS* = 5, and *KRAS* = 1) (Fig. 3*B*). Mutation co-occurrence analysis showed a significant association of *ATXN7* and *ARID2* mutations with *RAS* mutations (Fisher's exact P value = 0.005 and 0.021), respectively. One of the *RAS*-negative, *ATXN7*-positive tumors harbored an *NF1*-truncating mutation (p.R192X) (Fig. 3*B*), which functionally mimics RAS-activating mutations (Tables S1 and S2).

In addition to human thyroid tumor samples, we analyzed 41 authenticated human thyroid cancer cell lines (30) for gCIS mutations. As expected, the mutation burden of SB genes was higher in cell lines (2.76 mutations per sample) than in advanced thyroid carcinomas (0.62 mutation per sample), in part because there were no paired germline controls and more SB genes were mutated (40/45, 88.9%). Importantly, *ATXN7* (5/41, 12.2%), *ARID2* (5/41, 12.2%), and *PTEN* (5/41, 12.2%) were among the most recurrently mutated SB genes. *ATXN7* and *RAS* mutations co-occurred in thyroid cancer specimens and in cell lines (Fisher's exact P value = 0.003) (Dataset S4).

ATXN7 Mutants Cooperate with Oncogenic HRAS in Transformation. ATXN7 encodes for ataxin-7, a subunit of the highly conserved SAGA chromatin-modifying complex, which regulates gene expression by modulating histone acetylation and ubiquitination. Germline mutations leading to the expansion of a polyglutamine tract within this gene confer a predisposition to spinocerebellar ataxia 7 (31). The somatic mutations of *ATXN7* in PDTCs and ATCs also clustered in this region of the gene (Fig. 4*A*). Expression of canonical (92Q) and thyroid cancer-derived (Q30R, Q39P, P40Q, and 42-43insP) Flag-tagged mutants of ataxin-7 increased proliferation of Nthy-ori cells, an SV40 large T-immortalized human follicular thyroid line, in the context of $HRAS^{G12V}$ expression but not in WT cells (Fig. 4*C*).

Discussion

The goal of this transposon-based mutagenesis screen was to identify genes that cooperate with oncogenic Hras to induce thyroid cancer, because endogenous expression of mutant Hras is insufficient to drive thyroid tumorigenesis in the mouse. In humans, *RAS* mutations are found along the entire spectrum of the disease but are markedly enriched in PDTC and ATC (3). The histology of the SB-driven thyroid tumors showed a remarkable resemblance to *RAS*-mutant human PDTCs, which are characterized by encapsulation, the presence of necrosis, a high mitotic rate, and invasiveness. Interestingly, transposon insertional mutagenesis induced mouse tumors only in the context of $Hras^{G12V}$ homozygosity. Several studies point to the requirement of amplification of the mutant *Hras* allele or loss of the wild copy for transformation (5, 20, 21). The inability of the SB transposon system to overcome this constraint lends further credence to the importance of this step for tumor initiation.

The majority of CIS genes identified in previous studies of SB mutagenesis were putative tumor suppressors that were disrupted by transposon-integration events (32, 33), and the same was true in our series. We reach this conclusion because of the random orientation of transposon insertions in the gCIS, their location within their respective genes, and the known properties of many of the affected genes. Oncogene activation by SB transposons is dependent on their engineered promoter elements, which may have distinct activity in different cell lineages (34). The T2/Onc2 SB transposon used here contains the 5′ LTR of the murine stem cell virus (MSCV), a promoter that is primarily active in cells of the hematopoietic system. Accordingly, T2/Onc2 mice harboring a ubiquitously expressed Cre-activatable transposase allele primarily developed T-cell lymphomas and to a lesser extent medulloblastomas, intestinal and pituitary neoplasms (19). Nonetheless, we have evidence that the MSCV LTR is active in thyroid cells, because we were able to confirm expression of predicted Ankrd11 fusion transcripts in thyroid tumors harboring transposon-integration events in the sense orientation within the gene. Based on current evidence, the main gain-of-function mutations associated with human *RAS*-mutant PDTCs are in *EIF1AX*, a component of the translational preinitiation complex activated through recurrent point mutations

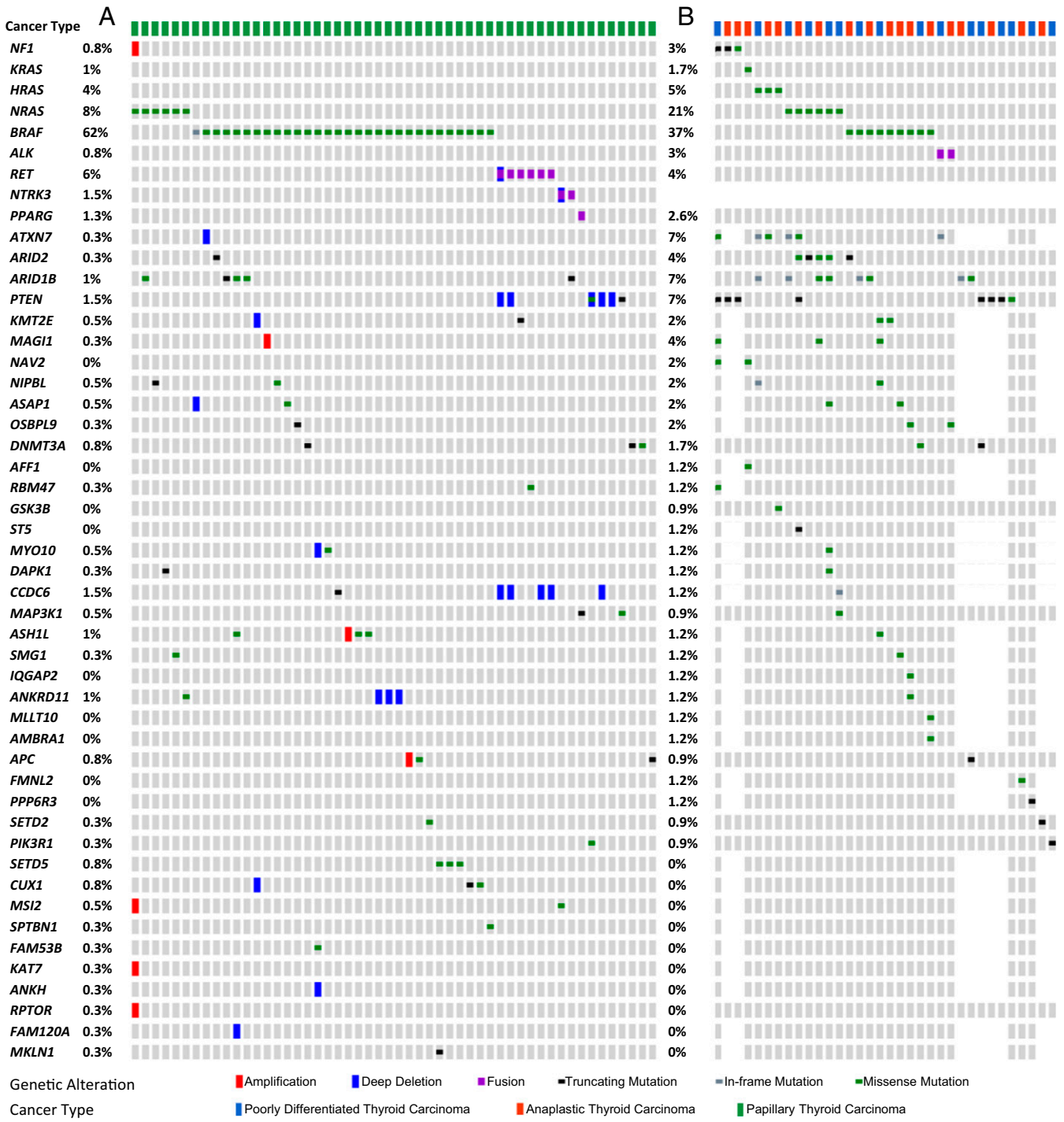

Fig. 3. Strong overlap of genes identified by SB transposon-integration events with genes mutated in human PDTC and ATC. (*A*) Oncoprint (mutations and copy number variations) of the SB gene list in well-differentiated papillary thyroid cancers from the THCA-TCGA series. Because most PTCs had a diploid genome, copy number variations are shown as possible indicators of tumorigenic events. (*B*) Oncoprint (mutations) of SB genes in a series of 84 PDTCs (blue) and 33 ATCs (red). The percent frequency of the listed mutations corresponds to those encountered in the entire series, whereas the color-coded boxes show mutations present in tumors that also harbored mutations of genes identified in the SB screen. The mutation type is color-coded according to the key at the bottom of the figure. Tumors with mutations in the major thyroid cancer-driver genes (*BRAF*, *NRAS HRAS*, *KRAS*, and *NF1* or fusions of *RET*, *ALK*, or *NTRK3*) are shown at the top of the oncoprint. Data were plotted with cBioPortal (www.cbioportal.org). Copy number variations, primarily affecting whole chromosome arms, are highly prevalent in advanced thyroid cancers (3) and are not represented in this oncoprint.

in specific hot spots, and in the *TERT* gene promoter, also associated with base substitutions at key residues (3). These mutations are not easily recapitulated by insertional mutagenesis with transposons. Hence, the predilection for transposon-driven inactivation of tumor suppressors is more likely caused by biological factors than by methodological biases.

The thyroid tumors in our series harbored a median of 10 gCIS events. Tumor development in SB mouse models has been proposed

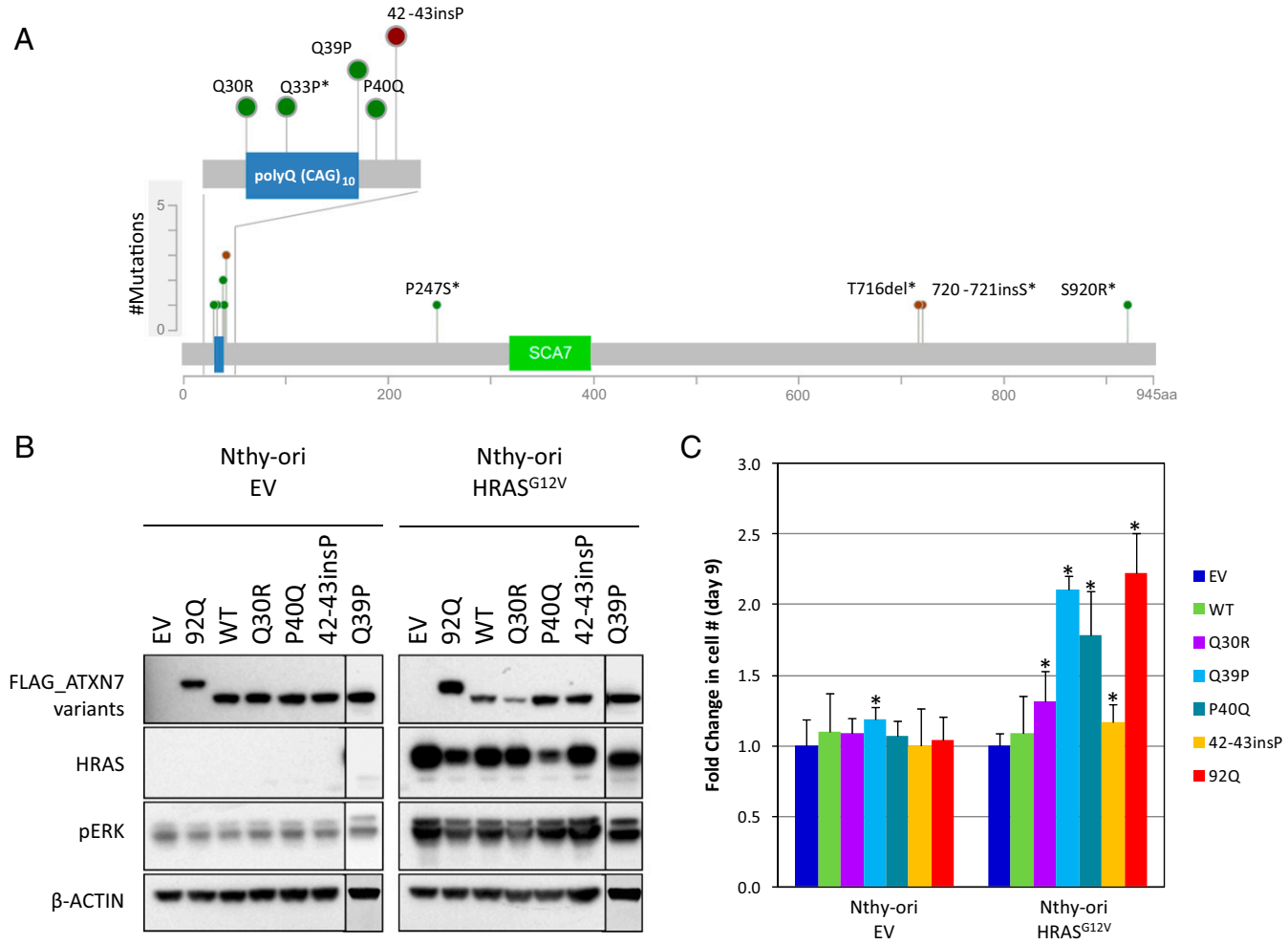

Fig. 4. ATXN7 mutants cooperate with oncogenic HRAS to induce thyroid cell proliferation. (*A*) Mutation map of *ATXN7* somatic mutations in PDTC and ATC. The mutations largely cluster in the N-terminal polyQ region of the gene. Green indicates missense mutations; brown indicates in-frame mutations. Asterisks denote mutations found only in thyroid cancer cell lines. (*B*) Western blots of Nthy-ori cells transfected with the indicated FLAG-ATXN7 variants with or without $HRAS^{G12V}$, hybridized with the indicated antibodies. The Q39P lanes were run in separate gels and were normalized to empty vector (EV) and FLAG-ATXN7 WT-expressing lysates. (*C*) Effects of expression of $HRAS^{G12V}$ and ATXN7 variants on proliferation of Nthy-ori cells 9 d after transfection (*$P < 0.05$; $n = 3$, unpaired *t* tests vs. empty vector).

to be caused by compound haploinsufficiency of many different tumor-suppressor genes that collectively contribute to tumor fitness. This conclusion is, at least for now, largely speculative. The recent development of a capture hybridization-based sequencing approach to identify transposon insertion sites from single cells may help clarify this question more definitively in the future (35).

Mutations of genes encoding effectors in the PI3K signaling pathway are comparatively infrequent in papillary thyroid cancers (about 2%) but are far more common in PDTC (11%) and ATC (39%) (3, 36). *PTEN* mutations in advanced thyroid cancers often co-occur with *NF1* or *RAS* mutations. *Pten* was among the most frequently targeted genes in our series. A minority of the tumors harbored two or more distinct transposon insertions in this gene, possibly representing convergent evolution events. Interestingly, in the samples available for IHC, Pten staining was absent and phospho-Akt was increased, regardless of whether there was one gCIS event or more within *Pten*. In one tumor with a single *Pten* gCIS there was a coexisting transposon-integration site within *Pik3r1*, which encodes for the p85 regulatory subunit of PI3K; the disruption of this subunit would be predicted to de-repress the catalytic activity of the enzyme. The association of RAS mutations with PTEN or other PI3K pathway mutations is not unique to thyroid cancers and has been reported in endometrial carcinomas, among others (37). The role of the Ras/Pten association in thyroid cancer pathogenesis has been demonstrated previously in Tpo-Cre/$Kras^{G12D}$/$Pten^{-/-}$ mice (7) and, as we now show here, in the context of the $Hras^{G12V}$ allele using the same genetic strategy.

Functional annotation of the SB hits identified serine/threonine protein kinases and phosphate metabolic processes as significantly affected groupings. As mentioned, PI3K signaling effectors (Pten, Gsk3beta, Raptor, Smg1, and Pik3R1) partially accounted for these functional categories. However, other kinases (i.e., Map3K1), adaptor molecules at the intersection of Mapkk and Pten (i.e., Magi1) (38), and phosphatases (i.e., Ppp3r1 and Ppp6r3) were also disrupted. A striking additional finding was the augmentation of MAPK pathway activity we saw in the tumors analyzed by IHC, which could not be accounted for by the gCIS identified in those specimens. These tumors phenocopy their human counterparts in this regard, because the transcriptional output of the MAPK pathway is markedly increased in advanced thyroid cancers (3).

Mutations of genes involved in regulating chromatin biology are markedly enriched in advanced forms of thyroid cancer compared with PTCs. More than one-third of ATCs have a mutation in one of

the genes that encode for components of the SWI/SNF chromatin remodeling complex (e.g., *ARID1A*, *ARID1B*, *ARID2*, *SMARCB1*, and *PBRM1*, among others), and a similar proportion harbor mutations of one of the histone methyltransferases (*KMT2D*, *KMT2A*, *KMT2C*, and *SETD2*) (3). Broadly, this class of genes was the most significant functional category altered by SB transposons in our model. SWI/SNF complexes consist of 12–15 subunits that, when assembled, hydrolyze ATP to mobilize nucleosomes and remodel chromatin. There is an evolutionarily conserved antagonism between SWI/SNF and polycomb repressor complexes (PRCs). The latter consist of at least five subunits, which, through the action of the histone lysine methyltransferases EZH1 or EZH2, catalyze the H3K27me3 repressive mark and lead to consequent transcriptional silencing of target genes. The canonical paradigm of this antagonistic relationship is in malignant rhabdoid tumors, where loss of SMARCB1 (also known as "SNF5"), a core component of SWI/SNF, leads to increased PRC2 complex activity and confers susceptibility to EZH2 inhibitors (39). However, loss-of-function mutations of PRC2 components have also been reported in several tumor types. Of particular interest is that in malignant peripheral nerve sheath tumors harboring NF1 mutations, the loss of PRC2 activity promotes tumorigenesis by amplifying a RAS-driven transcriptional program (40). The high prevalence of loss-of-function mutations of SWI/SNF components in thyroid cancers driven by RAS and other MAPK pathway effectors suggests that the PRC2-mediated effects on RAS-driven transcription may be context dependent.

Our screen also identified recurrent insertional events within *Atxn7*, a component of the SAGA complex. Germline mutations of *ATXN7* cause the autosomal dominant syndrome spinocerebellar ataxia 7. We identified somatic mutations of this gene in advanced thyroid cancers, particularly in those driven by RAS or NF1, in which it was present at high frequency (23% of cases). Relatively little is known about putative roles of *ATXN7* in cancer. A germline nonsynonymous SNP of *ATXN7* was one of the two most significant susceptibility variants conferring predisposition to breast cancer in a combined analysis of nine large genome-wide association studies (41). In addition, somatic fusions of *ATXN7* have been reported in colorectal and clear cell renal carcinomas (42, 43), but their functional consequence, if any, is unknown.

In conclusion, this study further buttresses the well-recognized role of PI3K pathway mutations in the progression of RAS-driven cancers. Most importantly, the results of this SB screen provide strong functional support for a key role of epigenetic modifiers that alter DNA methylation or chromatin structure in the progression of thyroid tumors driven by oncogenic RAS.

Materials and Methods

Generation of Hras^{G12V}/Tpo-Cre/T6113/Rosa LSL-SB Mice. FR-Hras^{G12V}/thyroid peroxidase-Cre (*Tpo-Cre*) mice were crossed with the 6113 transgenic line, which contains 358 copies of the transposon T2/Onc2 on chromosome 1 (19). These mice were then crossed with Rosa LSL-SB mice (11), in which a transposase vector has been inserted at the Rosa locus downstream of a lox-stop-lox element. The 6113 transgenic line and Rosa LSL-SB mice were kindly provided by Harold Varmus, Weill Cornell Medical College, New York, with permission from Neal Copeland, Houston Methodist Research Institute, Houston. This crossing resulted in mice expressing endogenous levels of Hras^{G12V} uniquely in thyroid follicular cells, as previously described (5, 6). Animal care and procedures were approved by the Memorial Sloan Kettering Cancer Center (MSKCC) Institutional Animal Care and Use Committee.

Laser-Capture Microdissection of Mouse Thyroid Nodules. Thyroids of quadruple-transgenic mice were followed up by bimonthly imaging with a Vevo 2100 micro-ultrasound system. Mice with abnormal thyroid nodular growth were killed, and thyroids were removed immediately using a dissecting microscope and embedded in optimum cutting temperature compound (OCT). OCT blocks were sectioned for H&E staining, and, after pathological assessment, serial 10-μm sections were mounted on microscopic slides and preserved at −80 °C until used for laser-capture microdissection. Frozen slides were stained with HistoGene (MDS Analytical Technologies), and the tumor area was defined by a pathologist and recovered with an Arcturus microdissection system (Life Technologies) at the MSKCC Pathology Core Facility.

DNA Digestion and Linker-Mediated PCR of Transposon-Insertion Sites. DNA from laser-capture microdissection samples and macroscopic tumors was isolated using the DNeasy Blood & Tissue Kit (QIAGEN). One hundred nanograms of DNA were digested with either AluI or NlaIII restriction enzymes, which on average have a cutting tag sequence every 300 bp in the gDNA. Because transposon ends also contain NlaIII and AluI sites, the digestion generates gDNA/transposon fragments at transposon-insertion loci. Next, double-stranded linkers were ligated to the ends of the fragments. A restriction digest with BamHI was performed to remove transposon fragments originated from the unmobilized transposons within the concatemer. A primary PCR with primers complementary to the adaptor and transposon end sequences specifically amplified the transposon-insertion sites. Nested PCR then was performed with modified barcoded primers to label each sample distinctly and to add sequences required for Illumina sequencing. Samples were pooled and sequenced on an Illumina HiSeq. 2000 lane (1 × 75 bp). The amount of gDNA was less than 100 ng in 12 of 26 tumors, so we performed whole-genome amplification (WGA) to generate the required amount of DNA. To test the reliability of WGA as a substrate for the detection of common transposon-insertion sites, we screened WGA and original DNA from sample 18804 by Hi-Seq. We identified 36 and 23 clonal sites in WGA 18804 and 18804, respectively, 17 of which were detected in both specimens (5 were WGA specific, and 19 were gDNA specific; however, all but one of these were identified as subclonal events in the other sample). Hence, although WGA sample analysis led to a partial depiction of the clonal insertion events, those that were detected were bona fide clonal insertion sites. This decrease in sensitivity likely accounts for the presence of gCIS events in only 3 of the 12 tumors subjected to WGA.

Identification of Candidate Genes. A list of clonal insertion sites in each sample was generated by including only those detected at a frequency higher than or equal to 5% of the most abundantly sequenced site within the sample. This conservative processing step was performed for reads detected from each side of the transposon, and the resultant clonal sites were used for downstream analyses. We used the gCIS calling method to determine RefSeq genes with a density of transposon insertion events higher than expected by chance (17). This computational method limits the analysis to transcribed regions of the genome and calculates the observed and expected number of transposon insertions. Briefly, the expected number of insertion events depends on the number of tumors analyzed, the total number of transposon insertions within each tumor, and the number of SB target sites (i.e., TA dinucleotides) within each RefSeq transcription unit, including 10 kb of promoter sequence. These values were used to perform a χ^2 test, yielding a P value for each RefSeq gene. A Bonferroni correction was applied to correct for multiple hypothesis testing.

Validation of Selected Transposon/Candidate Gene Fusions. Tissue from four fresh-frozen SB-induced tumors was isolated using TRIzol. RNA was reverse transcribed, and PCR was performed with primers bracketing the transposon splicing donor site and a site within the Ankrd11 cDNA flanking sequence. Primer sequences are shown in Fig. S1C.

Immunohistochemistry. Thyroid tissues were incubated overnight at 4 °C in 4% paraformaldehyde. The next day, tissues were washed twice with PBS for 30 min followed by a 30-min wash in 50% ethanol. The fixed tissues then were placed in 70% ethanol, paraffin embedded, and sectioned into 4-μm sections. H&E-stained slides were evaluated by a board-certified pathologist (R.G). Immunohistochemistry was performed on deparaffinized sections by incubating with antibodies to pERK (no. 4370; Cell Signaling Technology; 1 μg/mL), pAKT^{S473} (no. 4060; Cell Signaling; 1 μg/mL), and PTEN (no. 9188; Cell Signaling Technology; 4 μg/mL) at the MSKCC Molecular Cytology Core Facility as previously reported (44).

Massively Parallel Next-Generation Sequencing of Human Thyroid Tissues. Genomic sequencing of 84 PDTCs and 33 ATCs using the MSK-IMPACT cancer exome panel platform, which captures all exons and selected introns of 341 cancer genes (29), was recently reported (3). DNA obtained from a fresh-frozen or formalin-fixed, paraffin-embedded subset of these samples (62 PDTCs, 20 ATCs) was used to sequence the exomes of the 35 protein-coding gCIS genes not represented in MSK-IMPACT. Seventy-one of these samples had paired normal tissue, and the remaining 11 were filtered with pooled normal controls. Samples were sequenced to a depth of 500× using target hybridization with custom oligonucleotide probes derived from the NimbleGen 2.1M human exome array that were designed to capture all protein-coding exons of the 35 gCIS genes. Barcoded

sequence libraries were then generated, and captured exons were sequenced on an Illumina HiSeq. 2000 (2 × 75-bp reads) and analyzed as described (Tables S1 and S2) (45). All tissues were collected through the MSKCC Pathology and Tissue Bank with a protocol approved by the Institutional Review Board of the MSKCC.

Random Permutation Experiments for Assessment of Statistical Significance in the Subsets of SB Genes Mutated in Human Tumors. For the comparison with tumors from the TCGA-THCA dataset, we downloaded the study mutation annotation format (MAF) from the Genomics Data Commons Legacy Archive (https://portal.gdc.cancer.gov/legacy-archive). We randomly selected 45 gene identifiers from the full set of gene identifiers in the MAF and counted how many of these had a missense or truncating mutation in at least one sample. We repeated the previous step 1 million times to build a histogram of empirical frequencies. None of the 1 million random choices of genes that we evaluated resulted in 25 or more mutated genes, allowing us to estimate an empirical P value of $P < 1e\text{-}6$. For the comparison with the 82 human tumors analyzed with the augmented version of MSK-IMPACT, we proceeded in an analogous manner, leading to an observed empirical P value of $P = 0.0046$.

Plasmid Construction and Stable Cell Lines. pLenti-CMV-RasV12-Neo (HRAS-G12V) and pLenti-CMV-Neo-DEST (EV) constructs were from Addgene (w108-1 and 705-1, respectively). The Flag-tagged ATXN7 expression constructs (24Q and 92Q), kindly provided by Robert G. Roeder (Rockefeller University, New York), were subcloned into pLVX-Puro vector (Clontech). Mutant Flag-tagged ATXN7 mutant constructs (Q30R, Q39P, P40Q, and 42-43insP) were then generated by site-directed mutagenesis (QuikChange II XL kit; Agilent) using pLVX-Puro-F-ataxin-7-24Q as the template. Recombinant lentiviral particles were generated by transient transfection of HEK293T cells with Mission Lentiviral Packing Mix (Sigma-Aldrich) using Lipofectamine 2000 (Invitrogen). Nthy-ori cells were incubated with infectious particles in the presence of 10 ng/mL hexadimethrine bromide (Sigma) overnight. After recovery in complete medium for 24 h, cells were selected in medium with 300 μg/mL G418 (Geneticin) and 1 μg/mL puromycin.

Western Blotting. Western blotting of 9 μg of protein from the indicated cell lysates size-separated in 4–12% Bis-Tris NuPAGE Precast gels (Invitrogen) was performed following the manufacturer's instructions. The following primary antibodies (1:1,000) were added to the membrane with 5% nonfat dry milk in Tris-buffered saline and Tween 20: HRAS (Santa Cruz Biotechnology), pERK (Cell Signaling Technology), β-actin (Sigma-Aldrich), and FLAG antibody (Cell Signaling Technology). Secondary antibodies were HRP-conjugated goat anti-rabbit or goat anti-mouse antibodies (1:5,000; Santa Cruz). Bound antibodies were detected by chemiluminescence with the ECL detection system (GE Healthcare Biosciences).

Cell-Proliferation Assay. Nthy-ori cell lines expressing empty vector or $HRAS^{G12V}$ with or without FLAG-ATXN7 WT or mutants were plated at 0.5×10^5 cells per well in triplicate in six-well plates in RPMI medium with 10% FBS and 1% phosphate-buffered saline glucose. Cells were collected by trypsinization and were counted in a Vi-Cell series Cell Viability Analyzer (Beckman Coulter) at the indicated times.

ACKNOWLEDGMENTS. We thank Umeshkumar Bhanot of the Pathology Core and Sebastien Monet of the Molecular Cytology Core, Nikolaus Schultz for bioinformatic advice, and the Comparative Pathology and Animal Imaging Core Facilities of the Sloan Kettering Institute. This work was supported by NIH Grants CA50706, CA72597, P50-CA72012, and P30-CA008748, the Weiner Family Foundation, the David Linn family, the Frank D. Cohen Fund, and the Cole Family Fund. C.M.-C. was supported in part by a post-doctoral fellowship from the Fundación Asociación Española contra el Cáncer. F.S.-V. was supported by a grant from the Center for Metastasis Research of the Sloan Kettering Institute.

1. Namba H, Rubin SA, Fagin JA (1990) Point mutations of ras oncogenes are an early event in thyroid tumorigenesis. *Mol Endocrinol* 4:1474–1479.
2. Cancer Genome Atlas Research Network (2014) Integrated genomic characterization of papillary thyroid carcinoma. *Cell* 159:676–690.
3. Landa I, et al. (2016) Genomic and transcriptomic hallmarks of poorly differentiated and anaplastic thyroid cancers. *J Clin Invest* 126:1052–1066.
4. Schuhmacher AJ, et al. (2008) A mouse model for Costello syndrome reveals an Ang II-mediated hypertensive condition. *J Clin Invest* 118:2169–2179.
5. Chen X, et al. (2009) Endogenous expression of Hras(G12V) induces developmental defects and neoplasms with copy number imbalances of the oncogene. *Proc Natl Acad Sci USA* 106:7979–7984.
6. Chen X, Makarewicz JM, Knauf JA, Johnson LK, Fagin JA (2014) Transformation by Hras(G12V) is consistently associated with mutant allele copy gains and is reversed by farnesyl transferase inhibition. *Oncogene* 33:5442–5449.
7. Miller KA, et al. (2009) Oncogenic Kras requires simultaneous PI3K signaling to induce ERK activation and transform thyroid epithelial cells in vivo. *Cancer Res* 69:3689–3694.
8. Franco AT, et al. (2011) Thyrotrophin receptor signaling dependence of Braf-induced thyroid tumor initiation in mice. *Proc Natl Acad Sci USA* 108:1615–1620.
9. Kunstman JW, et al. (2015) Characterization of the mutational landscape of anaplastic thyroid cancer via whole-exome sequencing. *Hum Mol Genet* 24:2318–29.
10. Garcia-Rendueles ME, et al. (2015) NF2 loss promotes oncogenic RAS-induced thyroid cancers via YAP-dependent transactivation of RAS Proteins and sensitizes them to MEK inhibition. *Cancer Discov* 5:1178–1193.
11. Dupuy AJ, et al. (2009) A modified sleeping beauty transposon system that can be used to model a wide variety of human cancers in mice. *Cancer Res* 69:8150–8156.
12. Riordan JD, et al. (2013) Identification of rtl1, a retrotransposon-derived imprinted gene, as a novel driver of hepatocarcinogenesis. *PLoS Genet* 9:e1003441.
13. Rahrmann EP, et al. (2009) Identification of PDE4D as a proliferation promoting factor in prostate cancer using a Sleeping Beauty transposon-based somatic mutagenesis screen. *Cancer Res* 69:4388–4397.
14. Moriarity BS, et al. (2015) A Sleeping Beauty forward genetic screen identifies new genes and pathways driving osteosarcoma development and metastasis. *Nat Genet* 47:615–624.
15. Starr TK, et al. (2011) A Sleeping Beauty transposon-mediated screen identifies murine susceptibility genes for adenomatous polyposis coli (Apc)-dependent intestinal tumorigenesis. *Proc Natl Acad Sci USA* 108:5765–5770.
16. Heltemes-Harris LM, et al. (2016) Sleeping Beauty transposon screen identifies signaling modules that cooperate with STAT5 activation to induce B-cell acute lymphoblastic leukemia. *Oncogene* 35:3454–3464.
17. Brett BT, et al. (2011) Novel molecular and computational methods improve the accuracy of insertion site analysis in Sleeping Beauty-induced tumors. *PLoS One* 6:e24668.
18. Collier LS, Carlson CM, Ravimohan S, Dupuy AJ, Largaespada DA (2005) Cancer gene discovery in solid tumours using transposon-based somatic mutagenesis in the mouse. *Nature* 436:272–276.
19. Dupuy AJ, Akagi K, Largaespada DA, Copeland NG, Jenkins NA (2005) Mammalian mutagenesis using a highly mobile somatic Sleeping Beauty transposon system. *Nature* 436:221–226.
20. Bremner R, Balmain A (1990) Genetic changes in skin tumor progression: Correlation between presence of a mutant ras gene and loss of heterozygosity on mouse chromosome 7. *Cell* 61:407–417.
21. To MD, Rosario RD, Westcott PM, Banta KL, Balmain A (2013) Interactions between wild-type and mutant Ras genes in lung and skin carcinogenesis. *Oncogene* 32:4028–4033.
22. Volante M, et al. (2009) RAS mutations are the predominant molecular alteration in poorly differentiated thyroid carcinomas and bear prognostic impact. *J Clin Endocrinol Metab* 94:4735–4741.
23. Rivera M, et al. (2010) Encapsulated thyroid tumors of follicular cell origin with high grade features (high mitotic rate/tumor necrosis): A clinicopathologic and molecular study. *Hum Pathol* 41:172–180.
24. Huang W, Sherman BT, Lempicki RA (2009) Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 37:1–13.
25. Huang W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4:44–57.
26. Bartolomé RA, et al. (2015) IL13 receptor α2 signaling requires a scaffold protein, FAM120A, to activate the FAK and PI3K pathways in colon cancer metastasis. *Cancer Res* 75:2434–2444.
27. Gallagher D, et al. (2015) Ankrd11 is a chromatin regulator involved in autism that is essential for neural development. *Dev Cell* 32:31–42.
28. Futreal PA, et al. (2004) A census of human cancer genes. *Nat Rev Cancer* 4:177–183.
29. Cheng DT, et al. (2015) Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): A hybridization capture-based next-generation sequencing clinical assay for solid tumor molecular oncology. *J Mol Diagn* 17:251–264.
30. Schweppe RE, et al. (2008) Deoxyribonucleic acid profiling analysis of 40 human thyroid cancer cell lines reveals cross-contamination resulting in cell line redundancy and misidentification. *J Clin Endocrinol Metab* 93:4331–4341.
31. David G, et al. (1997) Cloning of the SCA7 gene reveals a highly unstable CAG repeat expansion. *Nat Genet* 17:65–70.
32. Mann KM, et al.; Australian Pancreatic Cancer Genome Initiative (2012) Sleeping Beauty mutagenesis reveals cooperating mutations and pathways in pancreatic adenocarcinoma. *Proc Natl Acad Sci USA* 109:5934–5941.
33. Mann MB, et al. (2015) Transposon mutagenesis identifies genetic drivers of Braf(V600E) melanoma. *Nat Genet* 47:486–495.
34. Howell VM (2012) Sleeping beauty–a mouse model for all cancers? *Cancer Lett* 317:1–8.
35. Mann KM, et al. (2016) Analyzing tumor heterogeneity and driver genes in single myeloid leukemia cells with SBCapSeq. *Nat Biotechnol* 34:962–972.
36. Fagin JA, Wells SA, Jr (2016) Biologic and Clinical Perspectives on Thyroid Cancer. *N Engl J Med* 375:1054–1067.

37. Cheung LW, et al. (2011) High frequency of PIK3R1 and PIK3R2 mutations in endometrial cancer elucidates a novel mechanism for regulation of PTEN protein stability. *Cancer Discov* 1:170–185.
38. Zmajkovicova K, et al. (2013) MEK1 is required for PTEN membrane recruitment, AKT regulation, and the maintenance of peripheral tolerance. *Mol Cell* 50:43–55.
39. Knutson SK, et al. (2013) Durable tumor regression in genetically altered malignant rhabdoid tumors by inhibition of methyltransferase EZH2. *Proc Natl Acad Sci USA* 110: 7922–7927.
40. De Raedt T, et al. (2014) PRC2 loss amplifies Ras-driven transcription and confers sensitivity to BRD4-based therapies. *Nature* 514:247–251.
41. Milne RL, et al.; GENICA Network; kConFab Investigators; Australian Ovarian Cancer Study Group; TNBCC (2014) Common non-synonymous SNPs associated with breast cancer susceptibility: Findings from the Breast Cancer Association Consortium. *Hum Mol Genet* 23:6096–6111.
42. Gotoh M, et al. (2014) Comprehensive exploration of novel chimeric transcripts in clear cell renal cell carcinomas using whole transcriptome analysis. *Genes Chromosomes Cancer* 53:1018–1032.
43. Kalvala A, et al. (2016) Rad51C-ATXN7 fusion gene expression in colorectal tumors. *Mol Cancer* 15:47.
44. Yarilin D, et al. (2015) Machine-based method for multiplex in situ molecular characterization of tissues by immunofluorescence detection. *Sci Rep* 5:9534.
45. Won HH, Scott SN, Brannon AR, Shah RH, Berger MF (2013) Detecting somatic alterations in tumor specimens by exon capture and massively parallel sequencing. *J Vis Exp*, e50710.