Supporting Information

Jendrzejewski et al. 10.1073/pnas.1205654109

SI Materials and Methods

Thyroid Tissue Samples. All samples included in this study were collected at The Ohio State University (OSU) as part of ongoing studies approved by the Institutional Review Board at OSU and were histologically of the papillary thyroid carcinoma type. All subjects gave written informed consent before participation. The samples were snap frozen in liquid nitrogen and stored at −80 °C. Total RNA was extracted using TRIzol solution according to the protocol of the manufacturer (Invitrogen) and stored at −80 °C.

Cell Lines. TPC-1 and BCPAP cells were a kind gift from Matthew D. Ringel and Motoyasu Saji (The Ohio State University, Columbus, OH). COS-7 and 293T cells were purchased from ATCC. TPC-1, COS-7, and 293T cells were cultured at 37 °C in 5% CO_2 in DMEM media (Gibco) supplemented with 10% FCS (Gibco). For the TPC-1 cells, the medium was also supplemented with 1× MEM nonessential amino acids (Gibco). BCPAP cells were cultured in RPMI medium supplemented with 10% FCS (Gibco).

In Silico Analysis. To search for the genes and ESTs in the vicinity of the rs944289, we used a genome browser (http://genome.ucsc.edu/cgi-bin/hgGateway, assembly February 2009, GRCh37/hg19). The potential transcription factor (TF) binding sites around the SNP were searched for by using the MATCH program (1). We retrieved two sequences 51 nucleotides long, one centered on the rs944289[C] allele and the other centered on the rs944289[T] allele. We scanned the retrieved sequences with MATCH in search of binding sites of known TF and calculated the Position Weight Matrix (PWM) score changes induced by the SNP (2).

Real-Time PCR Measurement of *PTCSC3* Expression. To investigate the expression of *PTCSC3* in pairs of unaffected thyroid and tumor tissue from patients diagnosed with PTC, we used SYBR Green real-time PCR (Applied Biosystems). Total RNA (1 μg) was reverse transcribed into cDNA by using the High Capacity Reverse Transcriptase kit (Applied Biosystems). The Master mix for each reaction of 10 μL total volume was prepared as follows: 2× SYBR Green Master mix 5μL, forward primer 200 nM (5′-TCAAACTCCAGGGCTTGAAC-3′), reverse primer 200 nM (5′-ATTACGGCTGGGTCTACCT-3′), cDNA 100 ng. The conditions for PCR performed on an ABI Prism 7900HT Sequence Detection System (Applied Biosystems) were as follows: 95 °C for 10 min to activate the polymerase followed by 40 cycles of denaturation (95 °C, 15 s) and annealing/elongation (60 °C, 1 min). For each plate, a dissociation curve was obtained to monitor any additional double stranded DNA. *GAPDH* was used as an internal control and the formula $2^{-\Delta Ct}$, where $\Delta Ct = Ct_{(GENE)} - Ct_{(GAPDH)}$ was used to calculate the relative mRNA level.

RLM-RACE. RLM-RACE kit (Ambion) was used to establish the full structure of CB987890 according to the protocol of the manufacturer. Both 10 μg and 1 μg total RNA extracted by TRIzol protocol from thyroid tissue (obtained from organ donors through Lifeline of Ohio; http://www.lifelineofohio.org) were used for 5′- and 3′-RLM-RACE, respectively. The CB987890 specific primers used in 3′-RLM-RACE were as follows: outer 5′-TGGGACCTGTTGTTTTTCTTG-3′; inner 5′-TCAAACTCCAGGGCTTGAAC-3′. For 5′-RLM-RACE the CB987890 specific primers were as follows: outer 5′-CATTTTGCCACCATGGTCTA-3′; inner 5′-TCTTCTGCTTGGCCTTTGAC-3′. Expand Long Template PCR system (Roche) was used for nested PCR reactions for both 5′- and 3′-RLM-RACE. The PCR conditions were as follows: 95 °C for 2 min followed by 10 cycles of denaturation (95 °C, 10 s) and annealing/elongation (60 °C, 30 s/68 °C, 3 min) followed by 20 cycles of denaturation (95 °C, 10 s) and annealing/elongation (60 °C, 30 s/68 °C, 3 min+ 20s for each successive cycle) with final elongation 68 °C, 5 min. The PCR RLM-RACE product was adenylated (10xPCR Buffer, 50mM dNTPs, 1 unit Taq polymerase, 72 °C, 10 min) and TOPO TA kit (Invitrogen) was used to clone nested amplicons into pCR4-TOPO vector according to the manufacturer's protocol. The clones were sequenced using M13 primers.

Northern Blot Analysis. Total RNA was isolated from 293T cells (negative control), 293T cells transfected with *PTCSC3* expression vector (positive control), and thyroid tissue obtained from organ donors through Lifeline of Ohio (http://www.lifelineofohio.org) by using TRIzol solution according to the protocol of the manufacturer (Invitrogen). A 20-μg quantity of RNA was separated on a 1% formaldehyde gel and transferred onto Hybond N membrane (GE Healthcare Life Sciences). The RNA was crosslinked (Stratalinker 1800 UV, Stratagene), and the membrane was hybridized at 42 °C in UltraHyb Ultrasensitive buffer (Ambion). A 183-bp DNA probe corresponding to 125–307 nucleotides of *PTCSC3* cDNA served as probe. Following overnight hybridization, the membrane was washed in 2xSSC/0.2% SDS at 42 °C.

Constructs. The cDNA of *PTCSC3* was amplified by PCR (forward primer: 5′-gtacggtaccCTCCTTCAGACTTCTCAGTACTC-3′; reverse primer: 5′-tcgactcgagATTGCTACTGTGAGCATAACCTAC-3′) and the amplicons were ligated between Kpn-I and Xho-I sites of the pcDNA3 vector (Invitrogen) to create the expression construct for *PTCSC3*. The construct was confirmed by Sanger sequencing. The promoter region (522 bp) of *PTCSC3* containing rs944289 was amplified by using forward (5′-atcaggtaccGGCAATTGAAGTTCCCAAAA-3′) and reverse (5′-atcactcgagGCCTCCAGACTTGGACTGAG-3′) primers from DNA of homozygous C and homozygous T patients. The amplicons were ligated between Kpn-I and Xho-I sites of the pGL4.24 vector (Promega) and the obtained constructs (C allele: rs944289-C and T allele: rs944289-T) were confirmed by sequencing. C/EBPα (p30 and p42) and C/EBPβ were cloned into pcDNA3-FLAG vector kindly provided by Daniel G. Tenen (Harvard Medical School, Boston, MA).

Luciferase Assay. COS-7 cells were seeded in quadruplicate into 24-well plates (4×10^4 cells per well) and after 24 h were transfected, as shown in Fig. 3, using Lipofectamine 2000 (Invitrogen). The empty pcDNA3-FLAG vector was used as carrier DNA to produce a constant amount of DNA in wells with no C/EBP constructs added, and phRL-TK-renilla reporter was used as transfection efficiency control. The cells were harvested 24 h after transfection using Passive Lysis Buffer (Promega). The *Firefly* and *Renilla* luciferase activity was measured using a Veritas Microplate Luminometer (Turner Biosystems). The promoter activity was determined as *Firefly*/*Renilla* ratio relative to the ratio obtained in cells transfected with the empty vector.

EMSA. 293T cells were transfected with an empty vector (pcDNA3-FLAG), p30 C/EBPα, p42 C/EBPα or C/EBPβ expression vectors. Next, the cells were harvested after 24 h and nuclear extract prepared. Briefly, the cells were washed once in PBS, resuspended in an equal volume of ice-cold hypotonic buffer A [10 mM Hepes-KOH (pH 7.9), 1.5 mM $MgCl_2$, 10 mM

KCl, 0.5 mM DTT, supplemented with Complete Mini Protease Inhibitor Mixture (Roche)], and incubated on ice for 15 min. Subsequently, the cells were lysed by 7–10 passages through a 26-gauge needle and nuclei were isolated by a 20-s centrifugation (14,000 rpm at 4 °C). Nuclear proteins were extracted by addition of 2/3 cell pellet volume of high-salt buffer C [20 mM Hepes-KOH (pH 7.9), 25% glycerol, 420 mM NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM DTT, supplemented with Complete Mini Protease Inhibitor Mixture (Roche)]. The nuclear extracts were recovered by centrifugation (14,000 rpm at 4 °C for 5 min). For EMSA the double-stranded oligonucleotide probe for the C allele (5′-GGAAAGATAGTCATTG*C*AGATTTGTAATA-3′) was labeled by filling in the ends with [α-P^{32}]-dCTP by Klenow. The nuclear extracts containing equal amounts of each C/EBP protein were incubated with 50,000 cpm radiolabeled oligonucleotide in a 25-μL reaction mixture containing 10 mM Hepes-KOH buffer (pH 7.9), 50 mM KCl, 2.5 mM $MgCl_2$, 1 mM DTT, 10% glycerol, 1 μg acetylated BSA, and 0.5 μg poly(dI-dC) on ice for 20 min. For the supershift, 1 μL of polyclonal anti-C/EBPα or anti-C/EBPβ antibody (Santa Cruz) was added to the binding reaction mixture. Binding reaction products were resolved on a 4% nondenaturing polyacrylamide gel containing 1× TBE (0.089 mM Tris, 0.089 mM boric acid, 0.002 mM EDTA) and electrophoresed at 150 V at 4 °C for 2 h.

***PTCSC3* cDNA Resequencing.** RNA from unaffected thyroid and tumor tissue (extracted using TRIzol solution by standard protocol) was DNase-1 treated (Ambion) and High Capacity Reverse transcriptase kit (Applied Biosystems) was used to obtain cDNA. After amplification by Expand Long Template PCR system (Roche) (forward primer: 5′-AAAAACTCCTTCAGACTTCTCAGT-3′; reverse primer: 5′-GTCCCAGTCCCAAAACTGAA-3′), the PCR amplicons were Sanger sequenced.

Microarray Hybridization Following Transfection with *PTCSC3*. TPC-1 cells (with no endogenous expression of *PTCSC3*) were plated in 12-well dish and transfected in quadruplicate with 50 ng of *PTCSC3* expression construct, and with 50 ng of empty vector (pcDNA3) as control using 2 μL of Lipofectamine 2000 reagent (Invitrogen). At 24 h after transfection, the cells were harvested and total RNA was extracted (TRIzol reagent, Invitrogen) and DNase-1 treated (Ambion). An Agilent BioAnalyzer 2100 (Agilent) was used to assess its integrity, and a High Capacity Reverse Transcriptase kit (Applied Biosystems) was applied to produce cDNA. To confirm successful transfection, RT-PCR was performed to detect *PTCSC3* expression (forward primer: 5′-TCAAACTCCAGGGCTTGAAC-3′; reverse primer: 5′-ATTACGGCTGGGTCTACCT-3′). SurePrint G3 Human Gene Expression 8 × 60k Arrays (AMADID 028004; Agilent Technologies) targeting 27,958 Entrez Gene RNAs and 7,419 long noncoding RNAs were used to assess gene expression. RNA samples were labeled with cyanine-3 (Cy3) using the Agilent Low Input One-Color Labeling kit (Agilent Technologies) and hybridized to the array for 17 h at 65 °C according to the manufacturer's protocol. Microarray slides were washed and then scanned with an Agilent G2505C Microarray Scanner. Images were analyzed with Feature Extraction 10.9.1 (Agilent Technologies) in one color gene expression mode. Median foreground intensities were obtained for each spot and imported into the mathematical software package R. The intensities were corrected for the scanner offset but not further background corrected. The dataset was filtered to remove positive control elements. Using the negative controls on the arrays, the background threshold was determined and all values less than this value were set to the threshold value. Finally, the data were normalized using the LIMMA microarray processing package in R (3).

Cell Growth Assay. TPC-1 and BCPAP cells were seeded into 100-mm plates (5×10^5 cells) and the next day transfected with 5 μg empty vector (pcDNA3) or *PTCSC3* construct using 12 μL Lipofectamine 2000 (Invitrogen). After 12 h, the cells were seeded in six replicates in 96-well plates (10^3 cells/well). Subsequently, 24 h, 48 h, and 72 h after transfection 10 μL of alamarBlue reagent (Invitrogen) was added to the cells, and after 1 h of incubation at 37 °C in 5% CO_2 the fluorescence intensity was measured by fluorescence microplate reader (SpectraMax M2, Molecular Devices).

Statistical Analysis. Raw gene expression data from Agilent microarrays were quantified and normalized, and replicated spots were averaged. All of the array expression data were log2 transformed before comparative analysis. The impact of *PTCSC3* compared with empty vector was evaluated by applying random variance model *t* test using BRB-Array Tools software (4). Network, functional, and canonical pathway analyses of differentially expressed gene expression data were performed using Ingenuity Pathway Analysis (IPA) software (Ingenuity Systems Inc). Real time PCR expression data and ratio data were compared by applying nonparametric Kruskal–Wallis analyses of variance tests and paired/nonpaired Wilcoxon rank sum tests. Luciferase and proliferation data were analyzed by applying *t* tests and are represented by mean ± SEM. All *P* values reported are two-sided.

1. Kel AE, et al. (2003) MATCH: A tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Res* 31:3576–3579.
2. Wingender E (2008) The TRANSFAC project as an example of framework technology that supports the analysis of genomic regulation. *Brief Bioinform* 9:326–332.
3. Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol*, 10.2202/1544-6115.1027.
4. Wright GW, Simon RM (2003) A random variance model for detection of differential gene expression in small microarray experiments. *Bioinformatics* 19:2448–2455.

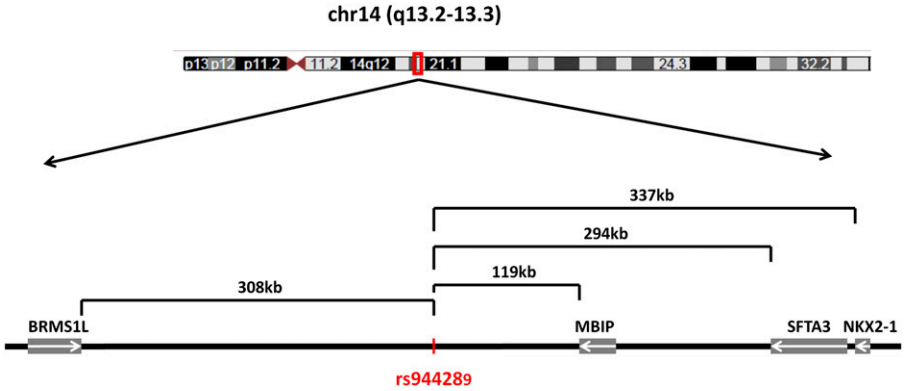

Fig. S1. Schematic view of the 14q13.3 region with the positions and distances between SNP rs944289 and the known neighboring genes.

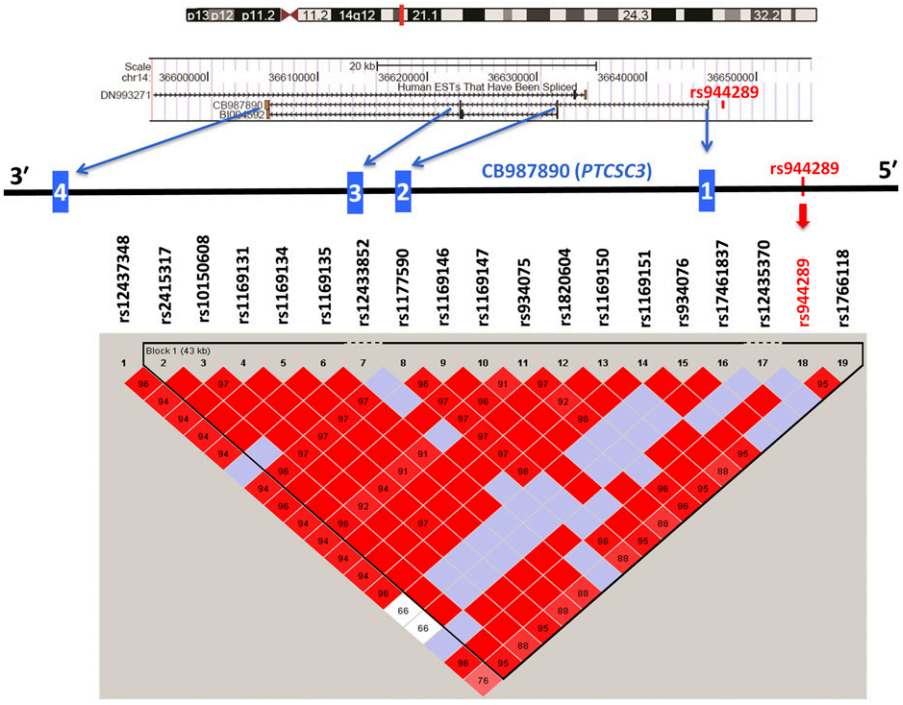

Fig. S2. Schematic view of rs944289 and CB987890 (*PTCSC3*) in a block of linkage disequilibrium (LD). (*Upper*) rs944289 (marked in red) and the candidate gene (CB987890/*PTCSC3*) located in the near vicinity of rs944289. (*Lower*) Schematic view of the LD structure of the region.

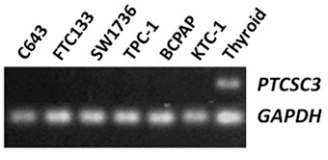

Fig. S3. *PTCSC3* expression assessed by RT-PCR in six thyroid carcinoma cell lines.

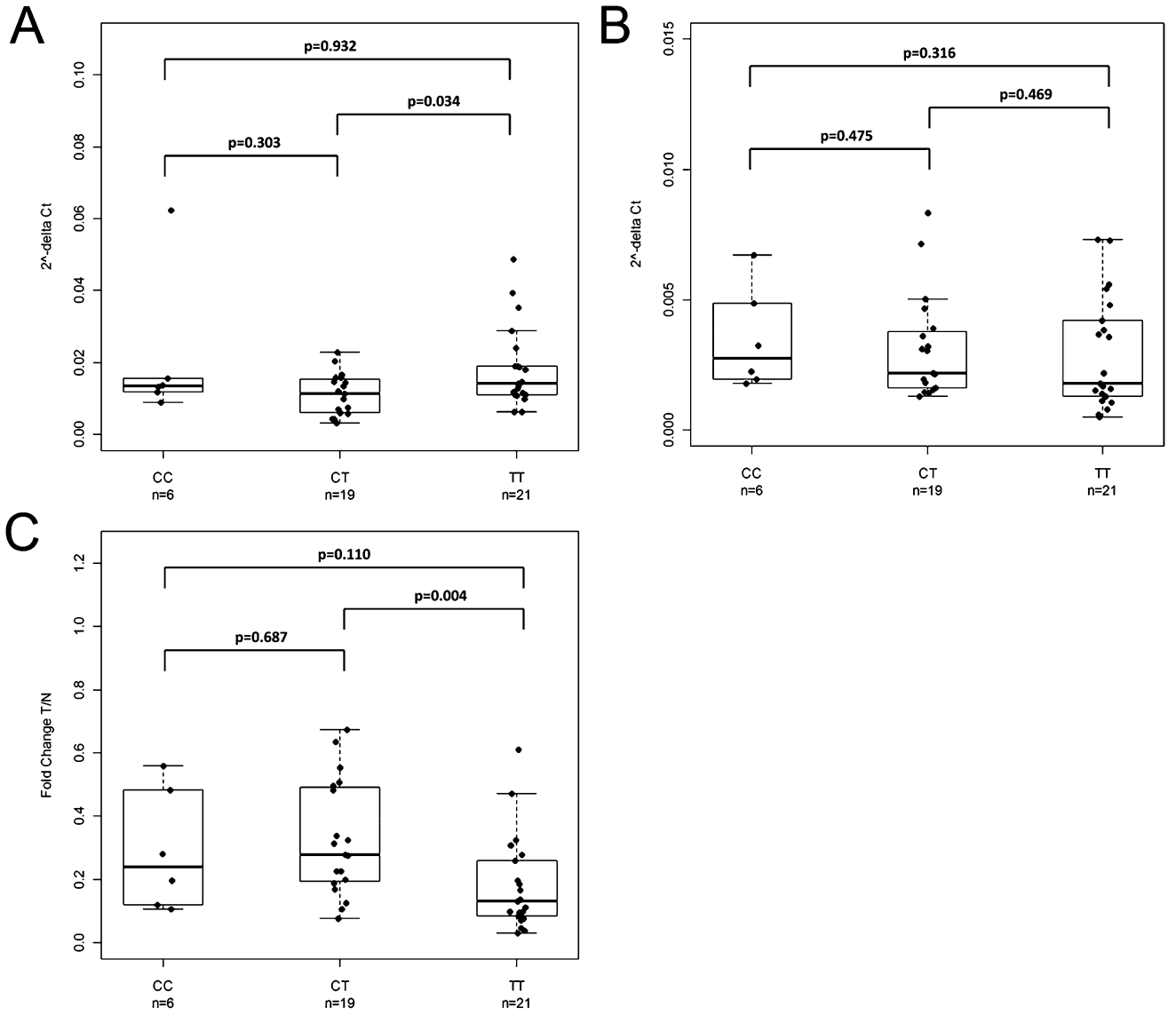

Fig. S4. *PTCSC3* expression measured by real-time PCR in 46 PTC tumor/unaffected tissue pairs. (*A*) *PTCSC3* expression in unaffected thyroid tissue. Homozygous [TT] patients show higher expression of *PTCSC3* in unaffected thyroid tissue compared with heterozygous [CT] subjects ($P = 0.034$) but not compared with homozygous [CC] patients ($P = 0.932$). (*B*) *PTCSC3* expression in thyroid tumor tissue. (C) Analysis of fold change of tumor/normal *PTCSC3* expression. Homozygous [TT] patients had stronger suppression of *PTCSC3* compared with heterozygotes ($P = 0.004$).

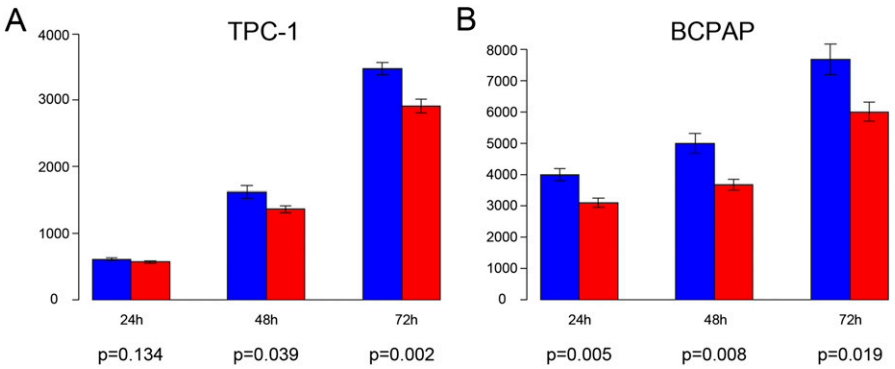

Fig. S5. Cell growth assay. (*A*) TPC-1 cells transiently transfected with *PTCSC3* construct show lower growth rate (red boxes) compared with cells transfected with empty vector (pcDNA3, blue boxes) ($P = 0.039$ and $P = 0.002$ at 48 h and 72 h, respectively). The *x* axis shows time points of fluorescence measurement; the *y* axis shows fluorescence intensity. The mean ± SEM of six independent experiments is shown at each time point. (*B*) BCPAP cells transiently transfected with *PTCSC3* construct show lower growth rate (red boxes) compared with cells transfected with empty vector (pcDNA3, blue boxes) ($P = 0.005$, $P = 0.008$, and $P = 0.019$ at 24 h, 48 h, and 72 h, respectively). The *x* axis shows time points of fluorescence measurement; the *y* axis shows fluorescence intensity. Mean ± SEM of six independent experiments is shown at each time point.

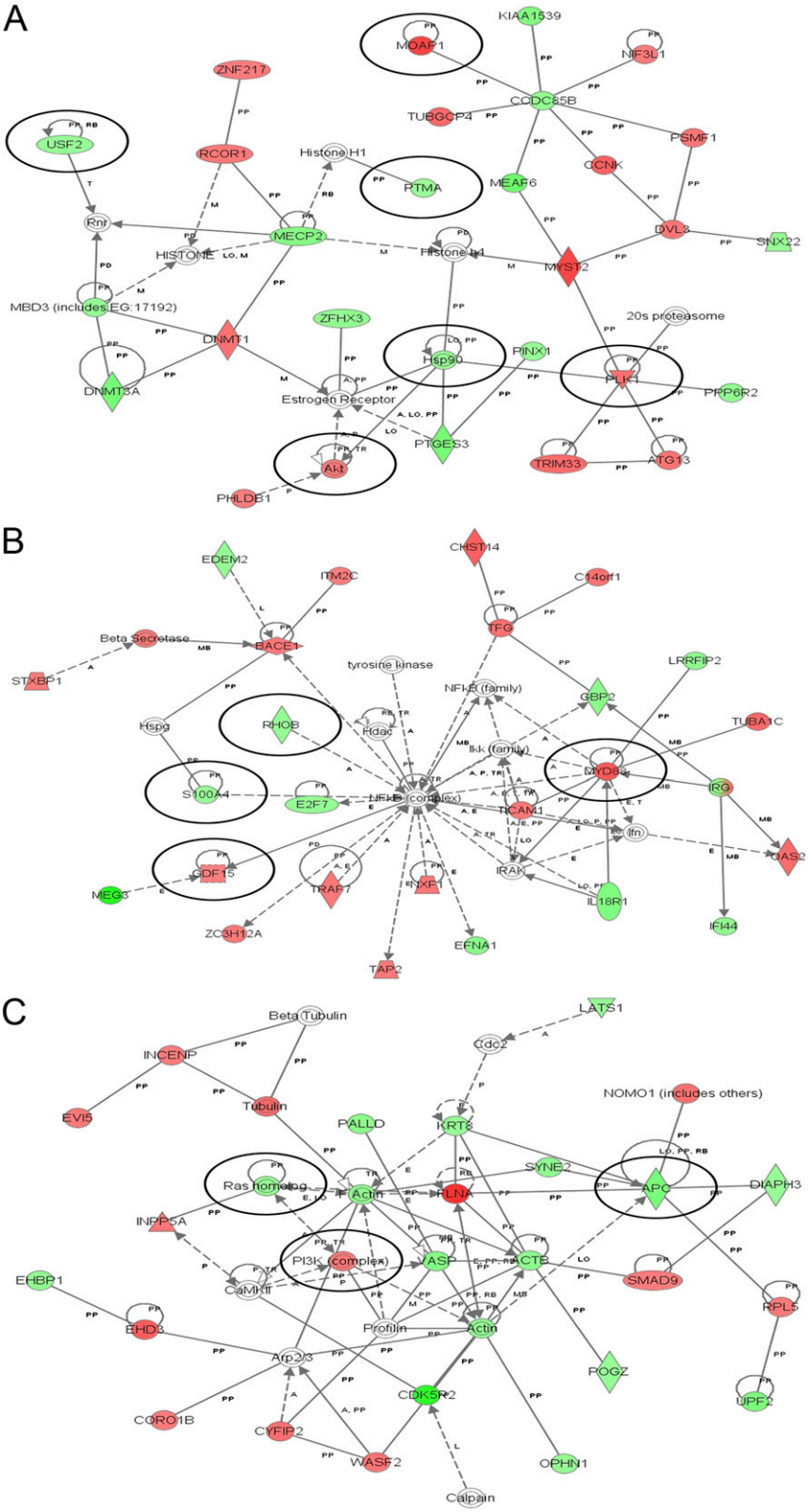

Fig. S6. IPA for gene interaction networks for transcripts affected by *PTCSC3*. Highest-scored networks are as follows: (*A*) network-1: DNA Replication, Recombination, and Repair, Gene Expression, Amino Acid Metabolism (IPA score = 45); (*B*) network-2: Cellular Movement, Tumor Morphology, Cell Death (IPA score = 40); (*C*) network-3: Cellular Assembly and Organization, Cellular Function and Maintenance, Tissue Morphology (IPA score = 39). Color indicates up-regulation (red) or down-regulation (green) of gene expression. Genes that have been linked with PTC in the literature are in black circles.

Table S1. Top five canonical pathways assigned to the genes affected by *PTCSC3* in the Agilent microarray expression assay generated by the Ingenuity Pathway Analysis (IPA) software

Canonical pathway	*P*	Ratio
Amyloid processing	3.8×10^{-4}	1.07×10^{-1}
DNA methylation and transcriptional repression signaling	5.62×10^{-4}	1.74×10^{-1}
Neuregulin signaling	9.54×10^{-4}	6.86×10^{-2}
Methionine metabolism	1.38×10^{-3}	5.13×10^{-2}
Beta gamma signaling	5.01×10^{-3}	5.13×10^{-2}

Significance of canonical pathways was tested by the Fisher exact test. Ratios reflect the number of molecules in a given pathway affected by *PTCSC3* divided by the total number of molecules that constitute that pathway.