SUPPLEMENTARY METHODS

Novel kinase fusion oncogenes in post-Chernobyl radiation-induced pediatric thyroid cancers

Julio C. Ricarte-Filho[1#], Sheng Li[2,3#], Maria E. R. Garcia-Rendueles[1], Cristina Montero-Conde[1], Francesca Voza[1], Jeffrey A. Knauf[1,4], Adriana Heguy[1], Agnes Viale[5], Tetyana Bogdanova[6], Geraldine A Thomas[7], Christopher E. Mason[2,3*], James A. Fagin[1,4*]

[1]Human Oncology and Pathogenesis Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10065
[2]Dept. of Physiology and Biophysics, Weill Cornell Medical College, New York, NY 10065 USA
[3]The HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational Biomedicine, Weill Cornell Medical College, New York, NY 10065 USA
[4]Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY 10065
[5]Genomics Core, Memorial Sloan-Kettering Cancer Center, New York, NY 10065
[6]Institute of Endocrinology and Metabolism, Kiev, Ukraine
[7]Department of Surgery and Cancer, Imperial College, Charing Cross Hospital, London, UK

Bioinformatic analysis of RNA sequencing data

Fusion detection: We trimmed the first 13 bases and last 13 bases of the paired-end RNA sequencing reads to 50 bases as recommended by Snowshoes (1). We then mapped the trimmed reads to the hg19 using BWA (version 0.5.9-r16)(2). We used Snowshoes v2.0 to detect fusion events. False positive filtering followed the Snowshoes user manual. Specifically, the minimum distance between two genes for fusion candidates is 5 kb; the minimal number of read pairs that supports each fusion gene candidate is 10; the maximum number of fusion isoforms between two fusion partners is 20. We found 210 fusions in total across all 12 samples. After filtering out those fusions that were less than 5kb apart on the same chromosome, we were left with 96 genes. Of these, 16 were somatic fusion genes/isoforms, which were found in tumor tissues but not in matched normal. We also looked for supporting evidence for reciprocal fusions (fusions in opposite directions) with at least 3 unique junction reads required to support the fusion call.

Detection of variants: The first 2 and last 3 bases of each read in the paired-end RNA sequencing fastQ file were trimmed to remove low quality and high error base calls using fastx toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). Sequences were then aligned using Tophat (version 1.4.1) (3) and duplicates were removed (i.e. clonal reads with the same start and end reads after mapping) using Picard version 1.76 (http://picard.sourceforge.net/). We then used the Samtools (version 0.1.18)(4) mpileup command to pileup reads. We called somatic SNVS and INDELs using Varscan 2 (5;6). The thresholds for somatic variant calls are: they should be observed on reads from both strands, with a minimum tumor variant allele frequency of 20% and a maximum normal variant allele frequency of 5%. The p-value threshold for calling variants was 0.05. The base quality threshold is 15 and the mapping quality threshold is 15. The minimum coverage was 8 on tumor and 6 on normal. Also, if 3 SNVs were observed within a 10 bps window, or if the SNVs were near INDELs, they were removed. We further annotated the variants and predicted their effect on amino acid changes using snpEff[10]. In the current study, we only focused on the variants with a high/moderate effect based on snpEff impact classifications, including non-synonymous, frame shift, start lost, stop lost, and stop gained variants[10]. The final list of SNVs/INDELs with high/moderate effect is listed in Supplementary Tables 5.

Gene expression: We aligned paired-end RNA sequencing reads to the hg19 genome using Tophat and hg19 refseq annotation. We then calculated gene expression using rmake (http://physiology.med.cornell.edu/faculty/mason/lab/r-make/) for read count value. The gene

expression value is normalized by edgeR[11-13] calcNormFactors function using Trimmed Mean of M component (TMM) Normalization[14]. It normalized the compositional difference between the libraries. We also applied cpm function in edgeR[11-13] to calculate the counts per million (CPM) gene expression value. In the calculation of CPM, the raw read counts were divided by library size (which was normalized by TMM[14]) and multiplying by one million.

Sample clustering: To cluster the samples by gene expression, we use log 2 fold-change of CPM adjusted by 0.01. The hierarchical clustering method is "ward" and the distance method is "euclidean". Analysis and clustering were performed in the R statistical and programming environment (http://www.r-project.org/). Gene set for MAPK and PPARG pathways were obtained from (7) and (8), respectively.

Bioinformatics of WGS

Sequence alignment: Paired-end reads from the whole genome sequencing were aligned to the hg19 human genome using Burrows-Wheeler Aligner (BWA) (version 0.7.5)[2]. PCR duplicates were removed using picard version 1.76 (http://picard.sourceforge.net). Alignments were further refined by The Genome Analysis Toolkit (GATK) pipeline (version 2.5-2) for local indel realignment and base quality recalibration (9;10). The mean ± standard deviation of the genome coverage was 7.22 ± 0.60.

Somatic SNVs and Indels: We identified somatic mutations using Varscan (version 2.3.2). We used eight quality thresholds for somatic variant calls, including: (1) variant observed on reads from both strands, (2) a minimum tumor variant allele frequency of 20% and (3) a maximum normal variant allele frequency of 5%, (4) a p-value threshold for calling variants of 0.05, (5) base quality threshold of 15, (6) a mapping quality threshold of 15, (7) minimum coverage of 6 on tumor and normal samples and (8) if 3 SNVs were observed within a 10 bps window, or if the SNVs were near INDELs, they were removed. We further annotated the variants and predicted their effect on amino acid changes using snpEff[10]. Mann-Whitney-Wilcoxon test was used to compare the number of SNVs between the radiation-exposed and the sporadic groups. Known somatic SNVs in patient samples T35 and T38 were not directly called by Varscan due to low coverage at the mutation locus, but we found read pairs (T35: 3; T38: 1) supporting the alternative allele in tumor but not matched normal samples.

Somatic chromosomal rearrangements: We identified somatic chromosomal rearrangement using BreakDancer (version 1.1.2) (11). The minimum number of read pairs required was 2. Only inter-

chromosomal rearrangement or intra-chromosomal rearrangements with distances greater than 50 kb were taken for further analysis. Structural variations were annotated using Refseq annotation (*www.ncbi.nlm.nih.gov/refseq/*) and any rearrangements coming from the same gene or from the similar gene symbols were removed. Mann-Whitney-Wilcoxon test was used to compare the number of rearrangements between the two groups. Known chromosomal rearrangement in patient T5 and T27 were not directly called by BreakDancer due to low coverage at the breakpoint, but we found read pairs (T5: 1; T27: 13) spanning the known fusion partners in tumor samples but not matched normal.

Reference List

1. Asmann,Y.W., Hossain,A., Necela,B.M., Middha,S., Kalari,K.R., Sun,Z., Chai,H.S., Williamson,D.W., Radisky,D., Schroth,G.P. et al 2011. A novel bioinformatics pipeline for identification and characterization of fusion transcripts in breast cancer and normal cell lines. *Nucleic Acids Res.* **39**:e100.

2. Li,H., and Durbin,R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics.* **26**:589-595.

3. Kim,D., and Salzberg,S.L. 2011. TopHat-Fusion: an algorithm for discovery of novel fusion transcripts. *Genome Biol.* **12**:R72.

4. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G., and Durbin,R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* **25**:2078-2079.

5. Koboldt,D.C., Zhang,Q., Larson,D.E., Shen,D., McLellan,M.D., Lin,L., Miller,C.A., Mardis,E.R., Ding,L., and Wilson,R.K. 2012. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22**:568-576.

6. Koboldt,D.C., Chen,K., Wylie,T., Larson,D.E., McLellan,M.D., Mardis,E.R., Weinstock,G.M., Wilson,R.K., and Ding,L. 2009. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics.* **25**:2283-2285.

7. Pratilas,C.A., Taylor,B.S., Ye,Q., Viale,A., Sander,C., Solit,D.B., and Rosen,N. 2009. (V600E)BRAF is associated with disabled feedback inhibition of RAF-MEK signaling and elevated transcriptional output of the pathway. *Proc. Natl. Acad. Sci. U. S. A.* **106**:4519-4524.

8. Giordano,T.J., Au,A.Y., Kuick,R., Thomas,D.G., Rhodes,D.R., Wilhelm,K.G., Jr., Vinco,M., Misek,D.E., Sanders,D., Zhu,Z. et al 2006. Delineation, functional validation, and bioinformatic evaluation of gene expression in thyroid follicular carcinomas with the PAX8-PPARG translocation. *Clin. Cancer Res.* **12**:1983-1993.

9. DePristo,M.A., Banks,E., Poplin,R., Garimella,K.V., Maguire,J.R., Hartl,C., Philippakis,A.A., del,A.G., Rivas,M.A., Hanna,M. et al 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**:491-498.

10. McKenna,A., Hanna,M., Banks,E., Sivachenko,A., Cibulskis,K., Kernytsky,A., Garimella,K., Altshuler,D., Gabriel,S., Daly,M. et al 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**:1297-1303.

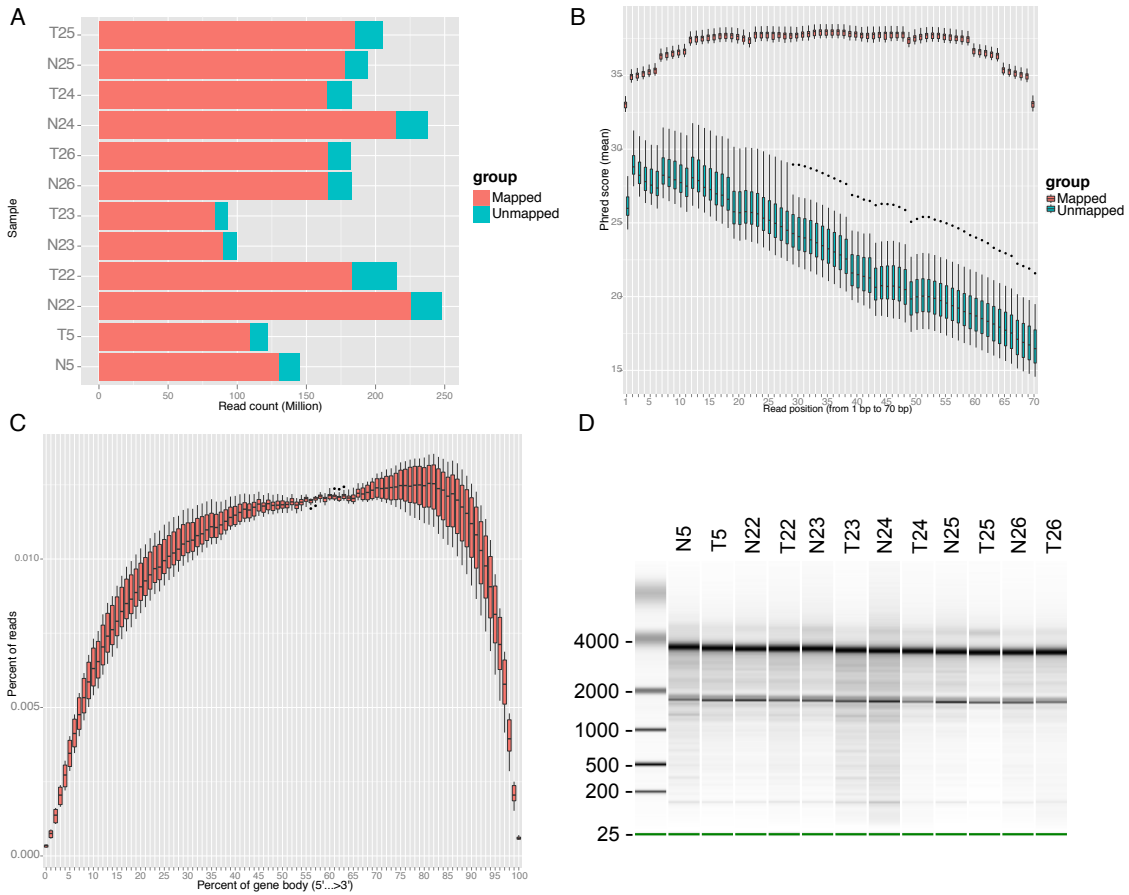
11. Chen,K., Wallis,J.W., McLellan,M.D., Larson,D.E., Kalicki,J.M., Pohl,C.S., McGrath,S.D., Wendl,M.C., Zhang,Q., Locke,D.P. et al 2009. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat. Methods.* **6**:677-681.

SUPPLEMENTAL DATA

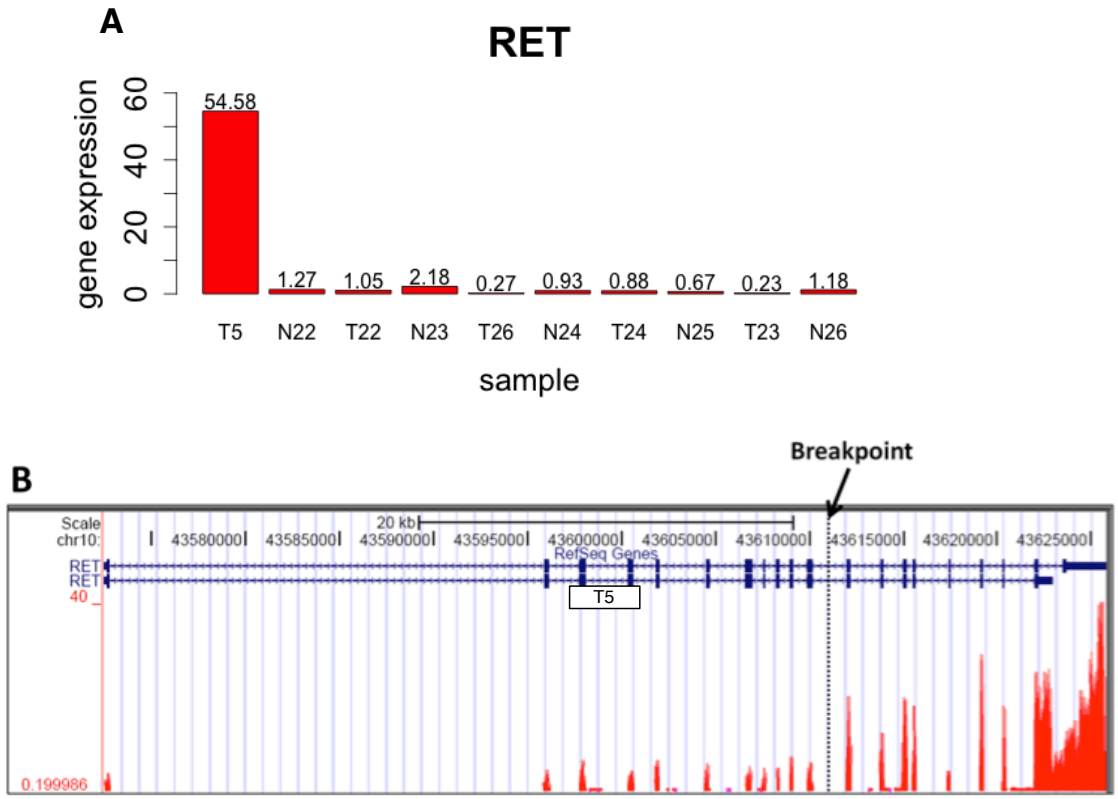
Novel kinase fusion oncogenes in post-Chernobyl radiation-induced pediatric thyroid cancers

Julio C. Ricarte-Filho[1#], Sheng Li[2,3#], Maria E. R. Garcia-Rendueles[1], Cristina Montero-Conde[1], Francesca Voza[1], Jeffrey A. Knauf[1,4], Adriana Heguy[1], Agnes Viale[5], Tetyana Bogdanova[6], Geraldine A. Thomas[7], Christopher E. Mason[2,3*], James A. Fagin[1,4*]

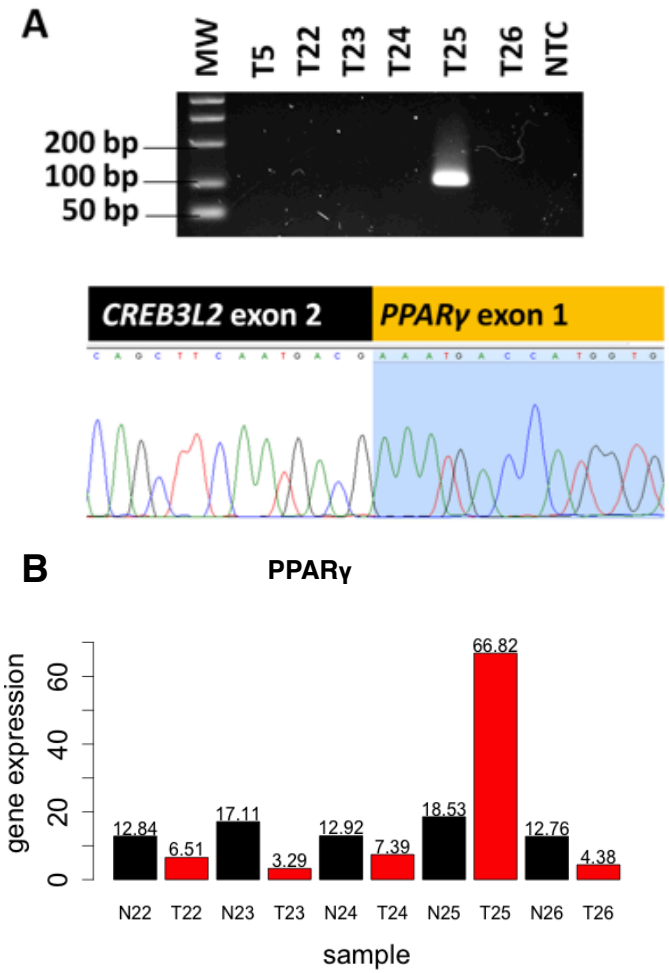
[1]Human Oncology and Pathogenesis Program, Memorial Sloan-Kettering Cancer Center (MSKCC), New York, NY 10065 USA.
[2]Dept. of Physiology and Biophysics, Weill Cornell Medical College (WCMC), New York, NY 10065 USA.
[3]The HRH Prince Alwaleed Bin Talal Bin Abdulaziz Al Saud Institute for Computational Biomedicine, WCMC, New York, NY 10065 USA.
[4]Department of Medicine, MSKCC, New York, NY 10065 USA.
[5]Genomics Core, MSKCC, New York, NY 10065 USA.
[6]Institute of Endocrinology and Metabolism, Kiev, Ukraine.
[7]Department of Surgery and Cancer, Imperial College, Charing Cross Hospital, London, UK.



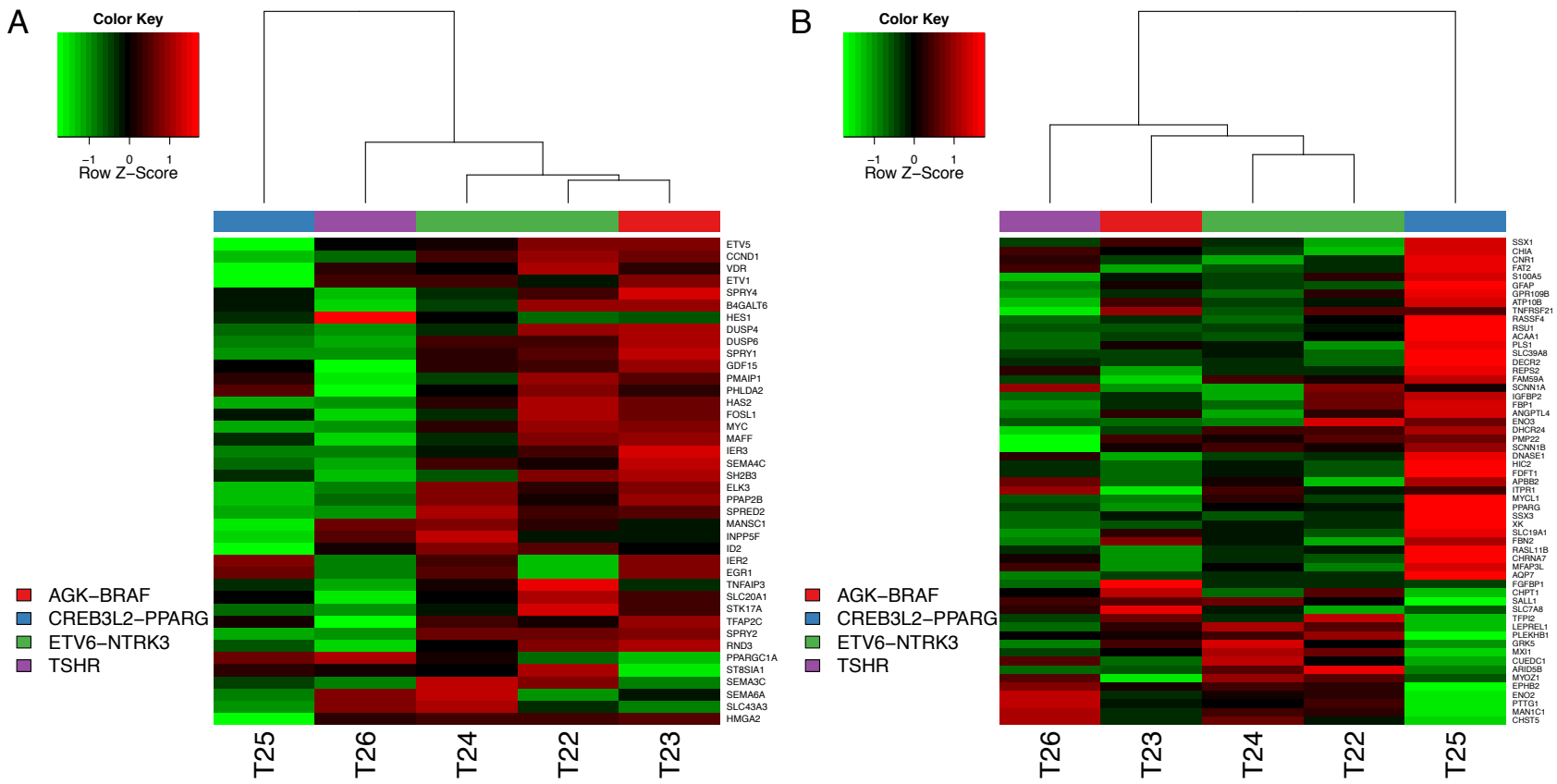
Supplementary Figure 1: RNAseq data quality control. A. Mapped and unmapped read count of all 12 RNAseq samples. **B.** Average phred quality score of mapped (red) and unmapped read (blue) for each base across all reads in 12 samples. Boxplot indicate the variance among samples. **C.** Gene body coverage over each percentage of all Refseq genes across all 12 RNAseq samples. Boxplot indicate the variance among samples. **D.** Bioanalyzer trace plot for RNA quality. The gene body coverage (C) indicates that there is a slight 3' enrichment for gene body coverage, which is consistent with the total RNA profile shown in the Bioanalyzer trace (D).



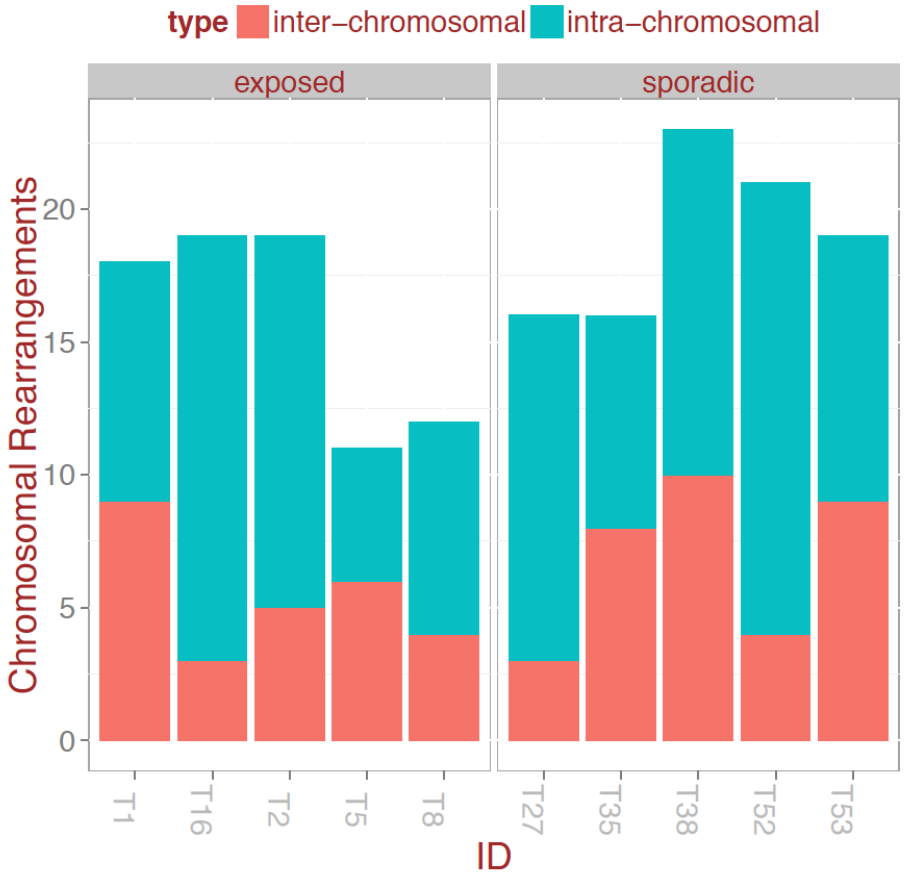
Supplementary Figure 2: RET gene expression in RNA Seq data of RET/PTC3 mutant thyroid cancer. A. Gene expression analysis confirmed *RET* mRNA to be exclusively enriched in the RET/PTC3 positive control sample. **B.** UCSC Genome Browser highlighting *RET* gene locus for sample T5. The breakpoint region between exons 11 and 12 of *RET* is shown. Increased FPKM values are seen in the 3' end of the gene (exons 12-20), which codes for the kinase domain of RET.



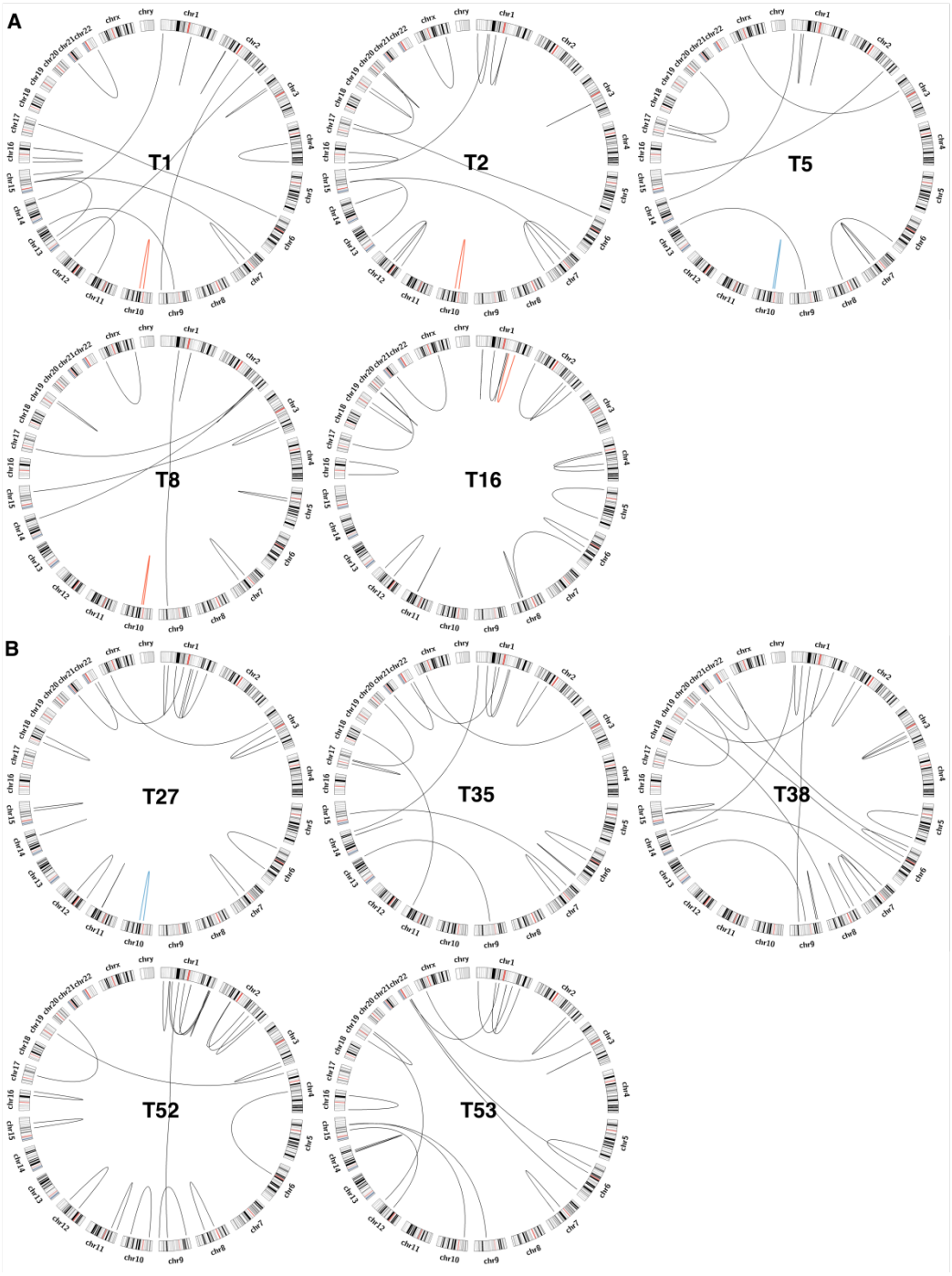
Supplementary Figure 3: **Identification and validation of CREB3L2/PPARG in a radiation-induced PTC. A.** Identification of CREB3L2/PPARG fusion by RT-PCR and confirmation by Sanger sequencing. **B.** *PPARG* expression (in CPM units) is increased exclusively in the tumor sample harboring the fusion.



Supplementary Figure 4: Heatmaps of gene expression log 2-fold change of tumor over normal in the 5 patient samples. A. Heatmap of MAPK regulated genes. Gene list was based on transcripts showing significant changes in expression in BRAF-mutant cell lines treated with or without MEK inhibitor[1]. **B.** Heatmap of PPARG-regulated genes, based on list derived from thyroid cancers harboring PAX8-PPARG fusions [2]. Z-score was calculated using log 2-fold change of CPM between tumors and normal and adjusted by 0.01.



Supplementary Figure 5: Total somatic chromosomal rearrangement counts in WGS of papillary thyroid cancers samples from radiation-exposed and sporadic cases. Inter-chromosomal rearrangements are shown in red, and intra-chromosomal rearrangements in blue. There is no significant difference in the frequency of these rearrangements between the two groups (p=0.288).



Supplementary Figure 6: Circos plots of somatic chromosomal rearrangements in thyroid cancers from radiation-exposed (A) and sporadic cases (B), as detected by low pass whole genome sequencing. Red lines represent the confirmed known somatic chromosomal rearrangement drivers called from the breakpoint detection pipeline (BreakDancer). Blue lines represent the driver rearrangements supported by paired end reads spanning previously identified drivers mutations. Black lines represent all other somatic chromosomal rearrangements as detected by BreakDancer. Known somatic chromosomal rearrangements: T1: RET/PTC1; T2: RET/PTC1; T5: RET/PTC3; T8: RET/PTC3; T16 NTRK1/TPR; T27: RET/PTC1.

Supplementary Table 1: Characteristics of radiation-exposed and sporadic pediatric thyroid cancers investigated in this study

A. Radiation-exposed

Sample ID	DOS	Age at op	Sex	Country	Oblast	Lesion size (cm)
T1	7-Jul-99	15.13	F	Ukraine	Kiev	4.2
T2	12-Jan-99	17.01	F	Ukraine	Zhytomyr	2.8
T3	19-Mar-04	18.79	F	Ukraine	Zhytomyr	1.7
T4	8-Jun-06	20.64	F	Ukraine	Kiev	3.5
T5	9-Apr-99	13.55	F	Ukraine	Kiev	2
T6	1-Nov-98	14.12	F	Ukraine	Kiev	1.5
T7	10-Dec-98	15.38	M	Ukraine	Kiev	1.3
T8	24-Dec-02	16.89	F	Ukraine	Rovno	2.3
T9	21-Feb-03	16.92	M	Ukraine	Kiev	5
T10	12-Sep-00	18.62	F	Ukraine	Kiev	3.5
T11	21-Dec-99	17.11	M	Ukraine	Kiev	2.5
T12	27-Jan-00	14.08	F	Ukraine	Kiev	2.5
T13	16-Feb-07	22.97	F	Ukraine	Kiev	3.4
T14	27-Nov-98	22.52	F	Ukraine	Kiev	1.3
T15	22-Jun-01	18.53	F	Ukraine	Kiev	2.8
T16	16-Feb-99	13.43	M	Ukraine	Chernigov	1.5
T17	21-Dec-01	19.99	F	Ukraine	Kiev	3.7
T18	5-Nov-03	19.61	F	Ukraine	Kiev	3.5
T19	17-Mar-04	20.94	F	Ukraine	Zhytomyr	3.8
T20	1-Apr-04	22.84	M	Ukraine	Chercassy	3
T21	26-Apr-02	16.01	F	Ukraine	Rovno	3.8
T22	23-Apr-99	15.69	F	Ukraine	Zhytomyr	5.9
T23	21-Sep-99	13.85	F	Ukraine	Kiev	2
T24	15-Nov-00	21.28	M	Ukraine	Kiev	3
T25	7-Jul-05	20.32	M	Ukraine	Chernigov	2.6
T26	4-Nov-05	22.08	F	Ukraine	Zhytomyr	3.2

B. Sporadic

Sample ID	DOS	Age at op	Sex	Country	Oblast	Lesion size (cm)
T27	21-Oct-09	17.17	F	Ukraine	Kiev	4.5
T28	11-Jan-07	16.62	M	Ukraine	Kiev	2.5
T29	20-Mar-07	15.15	F	Ukraine	Kiev	2.4
T30	8-Jun-07	18.22	F	Ukraine	Chernigov	2.8
T31	14-Nov-07	12.38	F	Ukraine	Zhytomyr	3
T32	5-Jun-02	14.15	F	Ukraine	Chercassy	7.5
T33	11-Dec-09	17.07	F	Ukraine	Kiev	2.5
T34	1-Jun-06	18.46	F	Ukraine	Kiev	3
T35	31-May-07	11.01	F	Ukraine	Chernigov	3.5
T36	15-Jun-07	15.56	F	Ukraine	Zhytomyr	1.7
T37	19-Jul-05	15.22	F	Ukraine	Sumy	4.5
T38	17-Apr-09	19.82	F	Ukraine	Chernigov	1.3
T39	14-May-09	17.53	F	Ukraine	Kiev	3.5
T40	21-May-09	13.06	F	Ukraine	Kiev	2.5
T41	23-Mar-99	11.56	F	Ukraine	Kiev	2.7
T42	4-Nov-08	19.32	M	Ukraine	Kiev	4.3
T43	12-Mar-03	13.97	F	Ukraine	Kiev	4.5
T44	17-Jun-04	17.09	F	Ukraine	Chercassy	1.7
T45	10-Feb-06	16.44	F	Ukraine	Zhytomyr	3
T46	21-Feb-06	16.07	F	Ukraine	Sumy	1.2
T47	30-Nov-07	18.01	F	Ukraine	Sumy	3.8
T48	11-Dec-07	18.33	F	Ukraine	Sumy	5.2
T49	13-May-08	19.62	F	Ukraine	Kiev	2.7
T50	25-Apr-00	8.79	F	Ukraine	Zhytomyr	1.6
T51	16-May-08	16.63	F	Ukraine	Kiev	3
T52	27-Feb-09	17.69	F	Ukraine	Kiev	5.5
T53	15-Sep-09	5.78	F	Ukraine	Rovno	3

DOS, date of surgery.

Supplementary Table 2: Complete list of somatic fusions identified by RNAseq in each tumor sample.

Sample	Alphabetical	Directional	Type	Exon1	Exon2	Distance	In_frame	Note
T23	AGK_BRAF	BRAF->AGK	intra-chr	chr7:AGK:141292945:141292985:+	chr7:BRAF:140500161:140500281:-	792664	yes	coding to coding fusion
T23	SIDT2_TAGLN	SIDT2->TAGLN	intra-chr	chr11:SIDT2:117066517:117066631:+	chr11:TAGLN:117073717:117073909:+	7086	no	Coding of SIDT2 to 5' of TAGLN
T23	VAMP5_VAMP8	VAMP8->VAMP5	intra-chr	chr2:VAMP5:85818847:85818985:+	chr2:VAMP8:85806131:85806290:+	12557	yes	coding to coding fusion
T26	CHCHD10_VPREB3	CHCHD10->VPREB3	intra-chr	chr22:CHCHD10:24108314:24108462:-	chr22:VPREB3:24094929:24095385:-	12929	yes	coding to coding fusion
T26	CHCHD10_VPREB3	CHCHD10->VPREB3	intra-chr	chr22:CHCHD10:24110020:24110141:-	chr22:VPREB3:24094929:24095385:-	14635	no	coding to coding fusion
T26	KIAA1984_TMEM141	TMEM141->KIAA1984	intra-chr	chr9:KIAA1984:139693553:139693675:+	chr9:TMEM141:139686702:139686810:+	6743	yes	coding to coding fusion
T26	SCNN1A_TNFRSF1A	SCNN1A->TNFRSF1A	intra-chr	chr12:SCNN1A:6457892:6457968:-	chr12:TNFRSF1A:6443256:6443410:-	14482	yes	coding to coding fusion
T24	ETV6_NTRK3	NTRK3->ETV6	inter-chr	chr12:ETV6:12022357:12022903:+	chr15:NTRK3:88669501:88669604:-	NA	yes	coding to coding fusion
T24	ETV6_NTRK3	ETV6->NTRK3	inter-chr	chr12:ETV6:12006360:12006495:+	chr15:NTRK3:88576087:88576276:-	NA	yes	coding to coding fusion
T24	IPO11_LRRC70	LRRC70->IPO11	intra-chr	chr5:IPO11:61887403:61887499:+	chr5:LRRC70:61874561:61874762:+	12641	NA	5'UTR of LRRC70 into the coding IPO11
T25	CREB3L2_PPARG	PPARG->CREB3L2	inter-chr	chr7:CREB3L2:137600582:137600758:-	chr3:PPARG:12329348:12329519:+	NA	NA	5'UTR of PPARG into the coding CREB3L2
T25	CREB3L2_PPARG	PPARG->CREB3L2	inter-chr	chr7:CREB3L2:137600582:137600758:-	chr3:PPARG:12353878:12353952:+	NA	NA	5'UTR of PPARG into the coding CREB3L2
T25	CREB3L2_PPARG	CREB3L2->PPARG	inter-chr	chr7:CREB3L2:137612895:137613112:-	chr3:PPARG:12421202:12421430:+	NA	yes	coding to coding fusion
T25	CTSD_IFITM10	CTSD->IFITM10	intra-chr	chr11:CTSD:1775032:1775131:-	chr11:IFITM10:1768896:1769349:-	5683	yes	coding to coding fusion
T25	GALT_IL11RA	GALT->IL11RA	intra-chr	chr9:GALT:34649406:34649561:+	chr9:IL11RA:34655214:34655314:+	5653	yes	coding to coding fusion
T25	SCNN1A_TNFRSF1A	SCNN1A->TNFRSF1A	intra-chr	chr12:SCNN1A:6457892:6457968:-	chr12:TNFRSF1A:6443256:6443410:-	14482	yes	coding to coding fusion

Supplementary Table 3: Somatic non-synonymous variants found by RNAseq in radiation-exposed pediatric thyroid cancers

Sample	Somatic	Chr	Pos	Gene	Aminoacid change /Indel effect	Tumor alternative allele frequency	Normal alternative allele frequency
T23	SNVs	chr12	94772742	CCDC41	R209Q	60	0
T26	SNVs	chr14	51238096	NIN	H358Y	36.36	0
T26	SNVs	chr14	81609676	TSHR	S425I	29.92	0
T26	SNVs	chr19	1484036	PCSK4	L387V	70	0
T26	SNVs	chr22	20094835	DGCR8	I647V	32.14	0
T26	SNVs	chr22	45821956	RIBC2	S263W	70	0
T26	SNVs	chr9	35853701	TMEM8B	R428H	28.07	0
T24	SNVs	chr1	1666175	SLC35E2	R229H	24.14	0
T24	SNVs	chr1	146049196	NBPF11,NBPF24	I392L,I338L	66.67	0
T24	SNVs	chr19	40383934	FCGBP	D3226N	26.92	0
T24	SNVs	chr19	40400342	FCGBP	R1916H	94.89	0
T24	SNVs	chr7	107834613	NRCAM	P526A	44.44	0
T25	SNVs	chr1	16377387	CLCNKB	H188Q	30	0
T25	SNVs	chr19	47910338	MEIS3	T314I	40	0
T25	SNVs	chr6	31846741	SLC44A4	R6L	43.14	0
T23	INDELs	chr10	32582030	EPC1	FRAME_SHIFT	36.36	0
T23	INDELs	chr14	60585130	C14orf135	FRAME_SHIFT	29.63	0
T23	INDELs	chr15	34445167	C15orf29	FRAME_SHIFT	33.33	0
T23	INDELs	chr16	68598462	ZFP90	FRAME_SHIFT	25	0
T23	INDELs	chr18	59739950	PIGN	FRAME_SHIFT	20	0
T23	INDELs	chr2	10051679	TAF1B	FRAME_SHIFT	22.73	0
T23	INDELs	chr3	32754743	CNOT10	FRAME_SHIFT	26.32	0
T26	INDELs	chr1	225142679	DNAH14	FRAME_SHIFT	21.74	0
T26	INDELs	chr15	52662410	MYO5A	FRAME_SHIFT	21.05	0
T26	INDELs	chr17	26925946	SPAG5	FRAME_SHIFT	21.62	0
T26	INDELs	chr19	58319467	ZNF552	FRAME_SHIFT	40	0
T26	INDELs	chr21	47773083	PCNT	FRAME_SHIFT	29.41	0
T26	INDELs	chr4	129869674	SCLT1	FRAME_SHIFT	27.78	0
T24	INDELs	chr11	64088132	PRDX5	FRAME_SHIFT	57.14	0
T24	INDELs	chr17	78302130	RNF213	FRAME_SHIFT	25.64	0
T24	INDELs	chr7	50514059	FIGNL1	FRAME_SHIFT	44.44	0
T25	INDELs	chr4	57220268	AASDH	FRAME_SHIFT	22.22	0

Supplementary Table 4: Somatic chromosomal rearrangements detected by low pass WGS in radiation-exposed and sporadic PTCs.

ID	Chr.1	Chr.2	Gene1	Gene2	Type	Coverage	Position 1	Position 2	Group
T2	chr1	chr1	KCNAB2	DOCK7	intra-chr.	2	6116348	63090946	exposed
T2	chr1	chr1	MTF2	PDE4DIP	intra-chr.	2	93552528	145022900	exposed
T2	chr1	chr1	NBPF9	PDE4DIP	intra-chr.	5	144680304	144954528	exposed
T2	chr1	chr15	YIPF1	MEF2A	inter-chr.	3	54324429	100112969	exposed
T2	chr10	chr10	RET	CCDC6	intra-chr.	7	43610671	61657290	exposed
T2	chr12	chr12	GXYLT1	MYO1A	intra-chr.	2	42533100	57438296	exposed
T2	chr12	chr12	RIMKLB	WSCD2	intra-chr.	2	8901787	108616919	exposed
T2	chr13	chr15	ENOX1	TYRO3	inter-chr.	5	44069573	41871244	exposed
T2	chr14	chr14	RBM25	PSEN1	intra-chr.	7	73552733	73608533	exposed
T2	chr16	chr16	LUC7L	ZNF423	intra-chr.	2	265617	49838363	exposed
T1	chr1	chr1	NBPF9	PDE4DIP	intra-chr.	5	144680292	144954536	exposed
T1	chr1	chr14	SLC2A5	BTBD7	inter-chr.	13	9121960	93712524	exposed
T1	chr10	chr10	RET	CCDC6	intra-chr.	12	43611594	61615267	exposed
T1	chr11	chr13	DDX10	SKA3	inter-chr.	11	108586716	21727905	exposed
T1	chr13	chr15	ENOX1	TYRO3	inter-chr.	10	44069703	41869990	exposed
T1	chr15	chr15	TYRO3	ETFA	intra-chr.	2	41853839	76553315	exposed
T1	chr16	chr16	NAA60	PRKCB	intra-chr.	2	3528177	24015208	exposed
T1	chr16	chr16	WWP2	DDX19B	intra-chr.	2	69932549	70327171	exposed
T1	chr2	chr12	ERBB4	PPP1R12A	inter-chr.	3	212279972	80310032	exposed
T1	chr2	chr2	MTA3	LRPPRC	intra-chr.	3	42798222	44142849	exposed
T1	chr2	chr9	DPP10	SET	inter-chr.	8	116377141	131457063	exposed
T1	chr21	chr22	BAGE3,BAGE2,BAGE5,BAGE4	FBLN1	inter-chr.	7	11023095	45991601	exposed
T1	chr3	chr3	CRTAP	ANO10	intra-chr.	2	33176800	43512365	exposed
T1	chr4	chr4	FAM13A	ASB5	intra-chr.	2	89672634	177170739	exposed
T1	chr6	chr17	PHACTR1	PRKCA	inter-chr.	7	13191467	64637361	exposed
T1	chr7	chr15	CBX3	C15orf57	inter-chr.	6	26252774	40854284	exposed
T1	chr7	chr7	STAG3L4	LOC100289561,LOC100630923	intra-chr.	13	66769841	102018998	exposed
T1	chr9	chr13	LOC642236	COL4A1	inter-chr.	9	68434485	110874356	exposed
T2	chr17	chr20	CDC27	PCIF1	inter-chr.	9	45214025	44570372	exposed
T2	chr19	chr19	SLC35E1	PDCD2L	intra-chr.	2	16661902	34898133	exposed
T2	chr20	chr20	COMMD7	ARHGAP40	intra-chr.	2	31323626	37272017	exposed
T2	chr3	chr3	STXBP5L	EAF2	intra-chr.	8	120912795	121577505	exposed
T2	chr6	chr17	PHACTR1	PRKCA	inter-chr.	6	13191424	64637318	exposed
T2	chr7	chr15	CBX3	C15orf57	inter-chr.	15	26252905	40853340	exposed
T2	chr7	chr7	GNA12	PTPRN2	intra-chr.	2	2865759	158284085	exposed
T2	chr7	chr7	STAG3L4	LOC100289561,LOC100630923	intra-chr.	19	66769495	102019027	exposed
T2	chrX	chrX	BEND2	MST4	intra-chr.	2	18229512	131165028	exposed
T5	chr1	chr1	NBPF9	PDE4DIP	intra-chr.	5	144680248	144954538	exposed

ID	Chr.1	Chr.2	Gene1	Gene2	Type	Coverage	Position 1	Position 2	Group
T5	chr1	chr1	PTPRF	HEATR8,HEATR8-TTC4	intra-chr.	2	44020913	55174458	exposed
T5	chr1	chr14	SLC2A5	BTBD7	inter-chr.	14	9121847	93712560	exposed
T5	chr17	chr17	SPECC1	MIR548W,MARCH10	intra-chr.	2	20201194	60819385	exposed
T5	chr17	chr20	CDC27	PCIF1	inter-chr.	15	45213111	44570629	exposed
T5	chr2	chr15	ALS2CR11	ETFA	inter-chr.	4	202475269	76540787	exposed
T5	chr3	chrX	LOC646498	DMD	inter-chr.	4	49218278	32000611	exposed
T5	chr6	chr8	AKAP12	RAB2A	inter-chr.	7	151653644	61526247	exposed
T5	chr7	chr7	STAG3L4	LOC100289561,LOC100630923	intra-chr.	15	66769407	102019657	exposed
T5	chr7	chr7	STYXL1	NAPEPLD	intra-chr.	2	75659291	102778761	exposed
T5	chr9	chr13	LOC642236	COL4A1	inter-chr.	7	68434819	110874401	exposed
T8	chr1	chr1	NBPF9	PDE4DIP	intra-chr.	4	144680264	144954530	exposed
T8	chr1	chr9	COL24A1	C9orf156	inter-chr.	5	86399171	100675648	exposed
T8	chr10	chr10	RET	NCOA4	intra-chr.	3	43611793	51583480	exposed
T8	chr19	chr19	MEGF8	ZNF600	intra-chr.	2	42841628	53275510	exposed
T8	chr2	chr14	SPAG16	NRXN3	inter-chr.	7	214761655	79654026	exposed
T8	chr2	chr17	C2orf80	SPECC1	inter-chr.	5	209033628	19995215	exposed
T8	chr2	chr2	ABI2	PARD3B	intra-chr.	2	204242988	205801045	exposed
T8	chr3	chr15	IGSF11	THSD4	inter-chr.	12	118698224	71491835	exposed
T8	chr3	chr3	NEK11	PPM1L	intra-chr.	4	130817209	160765678	exposed
T8	chr5	chr5	ARL15	MAST4	intra-chr.	2	53595815	66452136	exposed
T8	chr7	chr7	STAG3L4	LOC100289561,LOC100630923	intra-chr.	16	66769416	102019056	exposed
T8	chrX	chrX	HDHD1	MPP1	intra-chr.	2	7031201	154024054	exposed
T16	chr1	chr1	CLCA1	PI4KB	intra-chr.	2	86962406	151272918	exposed
T16	chr1	chr1	EPB41	PTPRU	intra-chr.	2	29301481	29597218	exposed
T16	chr1	chr1	NBPF9	PDE4DIP	intra-chr.	3	144680334	144954555	exposed
T16	chr1	chr1	NTRK1	MIR548F1,TPR	intra-chr.	10	156844239	186319105	exposed
T16	chr11	chr11	DLG2	ME3	intra-chr.	5	85217640	86254799	exposed
T16	chr12	chr12	ATF7	BRAP	intra-chr.	2	53931529	112119128	exposed
T16	chr16	chr16	TMC5	PLCG2	intra-chr.	2	19481158	81896154	exposed
T16	chr17	chr20	CDC27	PCIF1	inter-chr.	10	45213111	44570389	exposed
T16	chr19	chr19	PBX4	TMEM91	intra-chr.	3	19720966	41876476	exposed
T16	chr2	chr2	KLF11	FBXO36	intra-chr.	2	10185180	230826223	exposed
T16	chr2	chr2	TLK1	MFSD6	intra-chr.	2	172048248	191297328	exposed
T16	chr20	chr20	ISM1	MACROD2	intra-chr.	2	13249500	15664897	exposed
T16	chr21	chrX	GRIK1	PCDH11X	inter-chr.	3	30942674	91828049	exposed
T16	chr4	chr4	LPHN3	NAA15	intra-chr.	2	62658800	140242635	exposed
T16	chr4	chr4	USP46	RRH	intra-chr.	2	53511493	110756281	exposed
T16	chr5	chr5	CTNND2	GLRA1	intra-chr.	2	11494566	151277551	exposed
T16	chr6	chr6	ELOVL5	MDN1	intra-chr.	2	53196119	90461029	exposed
T16	chr6	chr8	MAN1A1	SNX16	inter-chr.	6	119558632	82754446	exposed

ID	Chr.1	Chr.2	Gene1	Gene2	Type	Coverage	Position 1	Position 2	Group
T16	chr8	chr8	WWP1	STK3	intra-chr.	2	87421039	99646721	exposed
T27	chr1	chr1	ASH1L	RASAL2	intra-chr.	2	155440438	178225665	sporadic
T27	chr1	chr1	FPGT-TNNI3K,TNNI3K	DNAH14	intra-chr.	2	75001409	225187293	sporadic
T27	chr1	chr1	NBPF9	PDE4DIP	intra-chr.	5	144680272	144954552	sporadic
T27	chr1	chr1	SPATA21	ZMYM4	intra-chr.	2	16762241	35884235	sporadic
T27	chr1	chr22	C1orf194	UQCR10	inter-chr.	5	109651044	30163398	sporadic
T27	chr11	chr11	DLG2	ME3	intra-chr.	2	85217654	86255789	sporadic
T27	chr12	chr12	FGF6	FAM19A2	intra-chr.	2	4547557	62384132	sporadic
T27	chr14	chr14	RBM25	PSEN1	intra-chr.	11	73551716	73608539	sporadic
T27	chr15	chr15	CCPG1,DYX1C1-CCPG1	MIR548H4	intra-chr.	2	55672749	69120363	sporadic
T27	chr18	chr18	LAMA3	WDR7	intra-chr.	2	21496141	54375094	sporadic
T27	chr21	chr22	BAGE3,BAGE2,BAGE5,BAGE4	FBLN1	inter-chr.	12	11023289	45991524	sporadic
T27	chr3	chr3	CADM2	MIR548H2	intra-chr.	2	85225492	151488160	sporadic
T27	chr3	chr3	STXBP5L	EAF2	intra-chr.	3	120912667	121577417	sporadic
T27	chr3	chrX	LOC646498	DMD	inter-chr.	4	49218270	32001142	sporadic
T27	chr6	chr6	ADTRP	TPD52L1	intra-chr.	2	11736193	125484860	sporadic
T27	chr7	chr7	STAG3L4	LOC100289561,LOC100630923	intra-chr.	11	66769380	102018999	sporadic
T35	chr1	chr1	DIO1	DAP3	intra-chr.	2	54366748	155700868	sporadic
T35	chr1	chr1	EVI5	NUF2	intra-chr.	2	93169798	163311916	sporadic
T35	chr1	chr14	SLC2A5	BTBD7	inter-chr.	12	9121755	93712056	sporadic
T35	chr1	chr22	C1orf194	UQCR10	inter-chr.	14	109650981	30163409	sporadic
T35	chr11	chr18	ARHGAP20	CCDC102B	inter-chr.	6	110496245	66572342	sporadic
T35	chr14	chr14	RBM25	PSEN1	intra-chr.	13	73551948	73608554	sporadic
T35	chr17	chr17	SARM1	STAT5B	intra-chr.	2	26720578	40385021	sporadic
T35	chr17	chr20	CDC27	PCIF1	inter-chr.	10	45213105	44570629	sporadic
T35	chr2	chr2	ALK	STAMBP	intra-chr.	2	29559033	74080414	sporadic
T35	chr21	chr22	BAGE3,BAGE2,BAGE5,BAGE4	FBLN1	inter-chr.	7	11023257	45991515	sporadic
T35	chr3	chrX	LOC646498	DMD	inter-chr.	14	49218369	32000635	sporadic
T35	chr6	chr6	GPLD1	C6orf147,KHDC1	intra-chr.	2	24451958	74005755	sporadic
T35	chr6	chr6	SLC22A2	PDE10A	intra-chr.	2	160678483	166010292	sporadic
T35	chr7	chr15	CBX3	C15orf57	inter-chr.	9	26248158	40853733	sporadic
T35	chr7	chr7	STAG3L4	LOC100289561,LOC100630923	intra-chr.	11	66769451	102019629	sporadic
T35	chr9	chr13	LOC642236	COL4A1	inter-chr.	8	68434548	110874969	sporadic
T38	chr1	chr1	DNAJC16	AGBL4	intra-chr.	2	15889308	50133340	sporadic
T38	chr1	chr1	NBPF9	PDE4DIP	intra-chr.	7	144680265	144954510	sporadic
T38	chr1	chr14	SLC2A5	BTBD7	inter-chr.	11	9121864	93712567	sporadic
T38	chr1	chr19	CHI3L1	ZNF793	inter-chr.	6	203150354	38031019	sporadic
T38	chr1	chr9	COL24A1	C9orf156	inter-chr.	10	86399221	100675135	sporadic
T38	chr14	chr14	RBM25	PSEN1	intra-chr.	13	73552879	73608536	sporadic
T38	chr15	chr15	UBR1	PTPLAD1	intra-chr.	2	43243426	65835053	sporadic

ID	Chr.1	Chr.2	Gene1	Gene2	Type	Coverage	Position 1	Position 2	Group
T38	chr17	chr20	CDC27	PCIF1	inter-chr.	11	45213105	44570496	sporadic
T38	chr2	chr2	TTC27	CNRIP1	intra-chr.	2	32983383	68531298	sporadic
T38	chr21	chr22	BAGE3,BAGE2,BAGE5,BAGE4	FBLN1	inter-chr.	7	11023295	45991509	sporadic
T38	chr3	chr3	ARHGAP31	MCM2	intra-chr.	3	119075643	127322495	sporadic
T38	chr3	chr3	MIR548G,COL8A1	RNF13	intra-chr.	2	99468841	149536385	sporadic
T38	chr5	chr5	PDE4D	KCNIP1	intra-chr.	2	58753839	170033407	sporadic
T38	chr6	chr20	SLC35A1	TSHZ2	inter-chr.	18	88217383	52021445	sporadic
T38	chr6	chr22	ACOT13	SYN3	inter-chr.	9	24683764	32927902	sporadic
T38	chr6	chr6	FARS2	PRIM2	intra-chr.	2	5360281	57389783	sporadic
T38	chr7	chr15	CBX3	C15orf57	inter-chr.	11	26246109	40854225	sporadic
T38	chr7	chr7	GTF2IRD1P1	ZNF425	intra-chr.	2	66292534	148805570	sporadic
T38	chr7	chr7	STAG3L4	LOC100289561,LOC100630923	intra-chr.	21	66769767	102019047	sporadic
T38	chr8	chr19	C8orf37	CRTC1	inter-chr.	3	96265755	18835172	sporadic
T38	chr8	chr8	CSMD1	DOCK5	intra-chr.	3	2983198	25047191	sporadic
T38	chr9	chr13	LOC642236	COL4A1	inter-chr.	9	68434471	110874521	sporadic
T38	chr9	chr9	PTPRD	BNC2	intra-chr.	2	9306942	16513699	sporadic
T52	chr1	chr1	ACOT11	WDR64	intra-chr.	3	55019303	241854112	sporadic
T52	chr1	chr1	HIPK1	RGS7	intra-chr.	3	114480803	241362838	sporadic
T52	chr1	chr1	NBPF9	PDE4DIP	intra-chr.	6	144680319	144954549	sporadic
T52	chr1	chr1	PCNXL2	RYR2	intra-chr.	3	233323067	237486004	sporadic
T52	chr1	chr1	ST3GAL3	LGALS8	intra-chr.	2	44310458	236685528	sporadic
T52	chr1	chr1	VPS13D	MACF1	intra-chr.	2	12451980	39904125	sporadic
T52	chr1	chr9	COL24A1	C9orf156	inter-chr.	9	86399004	100675238	sporadic
T52	chr10	chr10	ITGA8	WBP1L	intra-chr.	2	15574536	104538081	sporadic
T52	chr11	chr11	LMO1	PDHX	intra-chr.	2	8255492	34953286	sporadic
T52	chr12	chr12	DIP2B	GIT2	intra-chr.	2	51136872	110416421	sporadic
T52	chr15	chr15	GALK2	C15orf60	intra-chr.	2	49453181	73778939	sporadic
T52	chr16	chr16	DDX19B,LOC100506083	TCF25	intra-chr.	2	70357898	89947056	sporadic
T52	chr17	chr20	CDC27	PCIF1	inter-chr.	15	45213907	44570629	sporadic
T52	chr2	chr2	CWC22	MRPL44	intra-chr.	2	180821390	224829327	sporadic
T52	chr2	chr2	PCBP1-AS1	NEB	intra-chr.	3	70289674	152429678	sporadic
T52	chr2	chr2	ZNF638	ERBB4	intra-chr.	2	71596284	212619005	sporadic
T52	chr3	chr3	LINC00578	TM4SF19-TCTEX1D2	intra-chr.	2	177337524	196047479	sporadic
T52	chr4	chr19	WHSC1	NPAS1	inter-chr.	15	1881542	47529708	sporadic
T52	chr4	chr6	FAM190A	EYS	inter-chr.	5	91966842	64665999	sporadic
T52	chr8	chr8	KIAA0146	SDC2	intra-chr.	2	48579626	97548766	sporadic
T52	chr9	chr9	PTPRD	EXD3	intra-chr.	2	9901131	140313126	sporadic
T53	chr1	chr1	FNBP1L	RCOR3	intra-chr.	2	93924032	211480440	sporadic
T53	chr1	chr1	NBPF9	PDE4DIP	intra-chr.	8	144680249	144954561	sporadic
T53	chr1	chr1	SPSB1	RC3H1	intra-chr.	2	9393955	173920308	sporadic

ID	Chr.1	Chr.2	Gene1	Gene2	Type	Coverage	Position 1	Position 2	Group
T53	chr1	chr22	C1orf194	UQCR10	inter-chr.	23	109651023	30163375	sporadic
T53	chr10	chr15	CACNB2	VWA9	inter-chr.	14	18603761	65873083	sporadic
T53	chr12	chr15	BAZ2A	FSIP1	inter-chr.	9	56989988	39994167	sporadic
T53	chr12	chr19	ANKRD13A	FAM83E	inter-chr.	6	110455585	49111873	sporadic
T53	chr14	chr14	DCAF5	NRXN3	intra-chr.	2	69595851	79889574	sporadic
T53	chr14	chr14	RBM25	PSEN1	intra-chr.	7	73551919	73608525	sporadic
T53	chr16	chr16	ZNF263	HYDIN	intra-chr.	2	3337587	71045674	sporadic
T53	chr19	chr19	MLLT1	ZNF544	intra-chr.	2	6266040	58744131	sporadic
T53	chr2	chr2	MFSD6	SPAG16	intra-chr.	3	191320652	215011418	sporadic
T53	chr3	chr3	STXBP5L	EAF2	intra-chr.	2	120912711	121577503	sporadic
T53	chr3	chrX	LOC646498	DMD	inter-chr.	12	49218432	32000683	sporadic
T53	chr5	chr6	ADAM19	PRIM2	inter-chr.	9	156917135	57480540	sporadic
T53	chr6	chr22	ACOT13	SYN3	inter-chr.	14	24683718	32927891	sporadic
T53	chr6	chr22	ESR1	CABIN1	inter-chr.	6	152221372	24461882	sporadic
T53	chr7	chr7	STAG3L4	LOC100289561,LOC100630923	intra-chr.	13	66769664	102019855	sporadic
T53	chr9	chr15	RASEF	KIAA0101	inter-chr.	9	85622198	64664314	sporadic

Chr: chromosome

Supplementary Table 5: Somatics SNV counts in WGS reads of exposed and sporadic PTC cases.

Sample	Group	SNV	SNV per MB (%)
T1	exposed	10	0.32302935
T2	exposed	14	0.41993815
T5	exposed	9	0.29072641
T8	exposed	9	0.29072641
T16	exposed	19	0.61375576
T27	sporadic	6	0.19381761
T35	sporadic	14	0.45224108
T38	sporadic	9	0.29072641
T52	sporadic	15	0.48454402
T53	sporadic	14	0.45224108

Supplementary Table 6: Somatic non-synonymous coding missense SNVs in WGS reads of exposed and sporadic PTC cases.

Sample	Chromosome	Position	Gene	Aminoacid	Tumor frequency	Normal frequency	Group
T1	chr4	190874256	FRG1	T98M	24%	0%	exposed
T1	chr6	29523676	UBD	C160S	63%	0%	exposed
T1	chr7	100677816	MUC17	G1040E	56%	0%	exposed
T1	chr10	135082346	ADAM8	F592L	57%	0%	exposed
T1	chr10	135490610	DUX4L2	A74V	43%	0%	exposed
T1	chr11	1027811	MUC6	V619M	80%	0%	exposed
T1	chr11	58170291	OR5B3	I198V	71%	0%	exposed
T1	chr14	22102439	OR10G2	R187P	67%	0%	exposed
T1	chr19	55367314	KIR3DL2	A299D	71%	0%	exposed
T1	chr20	55027815	CASS4	R528L	85%	0%	exposed
T16	chr1	16902884	NBPF1	N666S	29%	0%	exposed
T16	chr1	144886197	PDE4DIP	F1013I	53%	0%	exposed
T16	chr2	97637905	FAM178B	I99M	75%	0%	exposed
T16	chr2	133542574	NCKAP5	V604M	83%	0%	exposed
T16	chr3	41756965	ULK4	V851I	63%	0%	exposed
T16	chr6	69666684	BAI3	N503S	71%	0%	exposed
T16	chr6	89926962	GABRR1	H27R	56%	0%	exposed
T16	chr7	12269417	TMEM106B	T185S	60%	0%	exposed
T16	chr10	129901726	MKI67	S2433N	86%	0%	exposed
T16	chr14	20872881	TEP1	N307K	57%	0%	exposed
T16	chr16	31925986	ZNF267	F139S	55%	0%	exposed
T16	chr19	8398006	KANK3	E610K	75%	0%	exposed
T16	chr19	55144100	LILRB1	T283A	71%	0%	exposed
T16	chr20	2539387	TMC2	R123K	57%	0%	exposed
T16	chr20	47258763	PREX1	V1240I	100%	0%	exposed
T16	chr20	55108617	FAM209B	Q74K	67%	0%	exposed
T16	chr21	14982886	POTED	G113S	71%	0%	exposed
T16	chr22	19968971	ARVCF	P220L	56%	0%	exposed
T16	chr22	41223190	ST13	M297I	67%	0%	exposed
T2	chr3	39178808	TTC21A	R1086Q	75%	0%	exposed
T2	chr4	17660082	FAM184B	R643H	63%	0%	exposed
T2	chr7	101989035	SPDYE6	R280C	24%	0%	exposed
T2	chr7	142574913	TRPV6	C157R	50%	0%	exposed
T2	chr8	144893084	SCRIB	P422L	80%	0%	exposed
T2	chr10	7749183	ITIH2	S60L	86%	0%	exposed
T2	chr10	129905429	MKI67	V1199M	78%	0%	exposed
T2	chr11	55563336	OR5D14	Q102L	44%	0%	exposed
T2	chr13	100543573	CLYBL	G310E	63%	0%	exposed
T2	chr15	43020983	CDAN1	R891C	78%	0%	exposed
T2	chr16	320614	RGS11	R351W	71%	0%	exposed
T2	chr16	84228770	ADAD2	G235R	88%	0%	exposed
T2	chr19	59082368	MZF1	R130Q	71%	0%	exposed
T2	chr20	3686436	SIGLEC1	V221M	100%	0%	exposed
T5	chr1	26582091	CEP85	S213N	86%	0%	exposed
T5	chr1	207646923	CR2	R671H	64%	0%	exposed
T5	chr3	75787321	ZNF717	G485R	29%	0%	exposed
T5	chr3	195513598	MUC4	P1618L	50%	0%	exposed
T5	chr9	33798574	PRSS3	S239N	56%	0%	exposed
T5	chr9	90503451	FAM75E1	R1350H	70%	0%	exposed
T5	chr10	68040325	CTNNA3	S596N	75%	0%	exposed
T5	chr10	101912064	ERLIN1	I291V	80%	0%	exposed
T5	chr13	20277413	PSPC1	M492V	36%	0%	exposed
T8	chr1	12854105	PRAMEF1	E110G	40%	0%	exposed
T8	chr1	18023690	ARHGEF10L	I1180V	67%	0%	exposed
T8	chr8	11995062	USP17L2	A403V	67%	0%	exposed
T8	chr8	132982824	EFR3A	N365D	63%	0%	exposed
T8	chr8	145741702	RECQL4	E267D	71%	0%	exposed
T8	chr9	2039793	SMARCA2	Q228P	78%	0%	exposed
T8	chr14	24707479	GMPR2	G242D	50%	0%	exposed
T8	chr15	63111739	TLN2	F2266L	56%	0%	exposed
T8	chr22	32586926	RFPL2	S324G	63%	0%	exposed
T27	chr3	148562310	CPB1	D208N	78%	0%	sporadic
T27	chr3	160804167	B3GALNT1	D126N	67%	0%	sporadic
T27	chr6	30993533	MUC22	R109G	80%	0%	sporadic

Sample	Chromosome	Position	Gene	Aminoacid	Tumor frequency	Normal frequency	Group
T27	chr7	117188736	CFTR	N417K	83%	0%	sporadic
T27	chr18	61647069	SERPINB8	R68Q	71%	0%	sporadic
T27	chr19	41384675	CYP2A7	H223R	80%	0%	sporadic
T35	chr3	75715181	FRG2C	G280R	27%	0%	sporadic
T35	chr3	75786355	ZNF717	T807S	26%	0%	sporadic
T35	chr7	99696797	MCM7	N144S	63%	0%	sporadic
T35	chr7	100679760	MUC17	G1688E	57%	0%	sporadic
T35	chr7	102939136	PMPCB	S74Y	57%	0%	sporadic
T35	chr10	50025396	WDFY4	R1816Q	63%	0%	sporadic
T35	chr11	1018366	MUC6	A1479T	67%	0%	sporadic
T35	chr13	53217493	HNRNPA1L2	Y289C	71%	0%	sporadic
T35	chr14	45645715	FANCM	N1253S	75%	0%	sporadic
T35	chr15	22369290	OR4M2	M239V	62%	0%	sporadic
T35	chr16	68344363	SLC7A6OS	S116A	67%	0%	sporadic
T35	chr17	39983820	NT5C3L	A209V	71%	0%	sporadic
T35	chr19	49206985	FUT2	G258S	57%	0%	sporadic
T35	chr19	55358655	KIR2DS4	Y237C	38%	0%	sporadic
T38	chr1	55224131	PARS2	N235S	67%	0%	sporadic
T38	chr5	35641582	SPEF2	N71H	100%	0%	sporadic
T38	chr7	33134883	RP9	K210R	55%	0%	sporadic
T38	chr9	7170006	KDM4C	S1037N	88%	0%	sporadic
T38	chr12	7277236	RBP5	M115L	67%	0%	sporadic
T38	chr17	62492582	POLG2	A169T	67%	0%	sporadic
T38	chr19	15289863	NOTCH3	R1231C	71%	0%	sporadic
T38	chr19	55450746	NLRP7	A481T	57%	0%	sporadic
T38	chr20	61528271	DIDO1	A556T	67%	0%	sporadic
T52	chr1	11775178	C1orf187	T284S	75%	0%	sporadic
T52	chr1	156263940	C1orf85	I137V	50%	0%	sporadic
T52	chr1	158324425	CD1E	Q106R	63%	0%	sporadic
T52	chr2	242757820	NEU4	G313R	67%	0%	sporadic
T52	chr3	50222926	SEMA3F	L503M	67%	0%	sporadic
T52	chr3	195507166	MUC4	L3762P	55%	0%	sporadic
T52	chr4	190989189	DUX2	L95M	100%	0%	sporadic
T52	chr4	190989190	DUX2	L95R	100%	0%	sporadic
T52	chr5	150901601	FAT2	R3518H	80%	0%	sporadic
T52	chr6	100868779	SIM1	P352T	56%	0%	sporadic
T52	chr13	46124056	FAM194B	T540A	60%	0%	sporadic
T52	chr15	59981515	BNIP2	V42A	67%	0%	sporadic
T52	chr17	47039132	GIP	S103G	50%	0%	sporadic
T52	chr20	23017082	SSTR4	F321S	71%	0%	sporadic
T52	chr22	25425282	KIAA1671	K439R	63%	0%	sporadic
T53	chr1	145281543	NOTCH2NL	T158I	29%	0%	sporadic
T53	chr5	108672946	PJA2	A705T	60%	0%	sporadic
T53	chr5	140574103	PCDHB10	L660V	57%	0%	sporadic
T53	chr5	150696496	SLC36A2	A445V	100%	0%	sporadic
T53	chr7	34979905	DPY19L1	G502V	55%	0%	sporadic
T53	chr8	124663987	KLHL38	G394R	86%	0%	sporadic
T53	chr9	33385784	AQP7	E202D	54%	0%	sporadic
T53	chr9	33385786	AQP7	E202Q	46%	0%	sporadic
T53	chr13	109496813	MYO16	M385T	60%	0%	sporadic
T53	chr19	39108034	MAP4K1	R70W	67%	0%	sporadic
T53	chr19	47291174	SLC1A5	P17A	71%	0%	sporadic
T53	chr19	52520607	ZNF614	G82R	67%	0%	sporadic
T53	chr21	47777063	PCNT	G704E	88%	0%	sporadic
T53	chr22	19808769	GNB1L	R37H	100%	0%	sporadic

Supplementary Table 7: Somatic indels in WGS reads of exposed and sporadic PTC cases.

Sample	Chromosome	Position	Gene	Effect	Tumor frequency	Normal frequency	Group
T8	chr21	34003928	SYNJ1	CODON_INSERTION	67%	0%	exposed
T27	chr6	170871013	TBP	CODON_CHANGE_PLUS_CODON_INSERTION	63%	0%	sporadic
T35	chr3	122459290	HSPBAP1	CODON_INSERTION	60%	0%	sporadic
T38	chr6	170871013	TBP	CODON_CHANGE_PLUS_CODON_INSERTION	83%	0%	sporadic
T38	chr17	4689194	VMO1	SPLICE_SITE_DONOR	75%	0%	sporadic
T52	chr17	18874685	FAM83G	CODON_INSERTION	64%	0%	sporadic

Supplementary Table 8: Primers used to validate the novel genetic alterations found in radiation-exposed pediatric thyroid cancer

Fusion/Mutation		Forward (5'>3')	Reverse (5'>3')	Size (bp)
ETV6-NTRK3	cDNA	ACACACACAGCCGGAGGTCATAC	AGTGGGCTGGCTGAGTCCTCC	90
AGK-BRAF	cDNA	CTGCTGACCTGGGGAGGCCATT	TCATCTGCTGGTCGGAAGGGCTG	113
CREB3L2-PPARg	cDNA	CCAGTCGCCCTTCACCCACATT	TCCACGGAGCTGATCCCAAAGT	111
TSHRS425I	gDNA	GGGGACAGTGAAGACATGGTGTG	CAGAGGCGATGAGGAGCAGGT	247