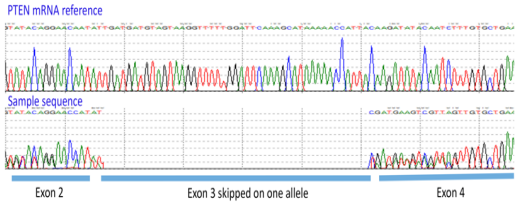
Supplemental Figures and Tables (Chen et al)

Supplemental Figure 1.

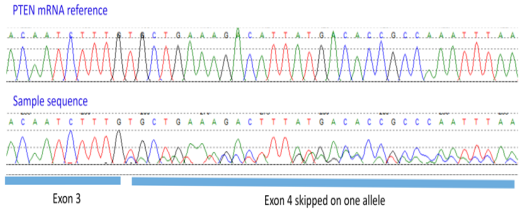
Sanger sequencing chromatograms showing skipping of exons and other aberrant splicing in *PTEN*[1] cDNA resulting from germline intronic variants. Nucleotide numbering uses +1 as the A of ATG translation initiation codon in the reference sequence, with the initiation codon as codon 1. **A**: Exon 3 skipping. **B**: Exon 4 skipping. **C**: Exon 5 skipping. **D**: Exon 6 skipping. **E**: Intron 2 mutation c.165-2A>G creates new splice site using -3 position A and mutant G resulting adding G in 5' end of exon3. **F**: *PTEN* intron 6 mutation c.635-1G>C uses cryptic splice site at r.642_643 causing mutant transcript missing 8 nucleotides at the 5' end of exon7. **G**: *PTEN* intron 6 mutation c.635-1G>C uses cryptic splice site at r.688_689 causing mutant transcript missing 54 nucleotides at the 5' end of exon7. **H**: Intron 7 mutations c.802-2A>T and c.802-2A>G use cryptic splice site at r.844_845 causing mutant transcript missing 43 nucleotides at the 5' end of exon 8.

[1] The *PTEN* reference sequence is NC_000010.11

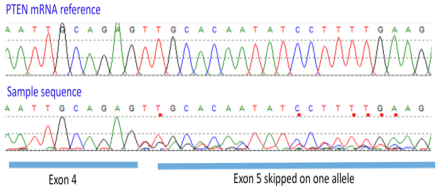
A Exon 3 skipping



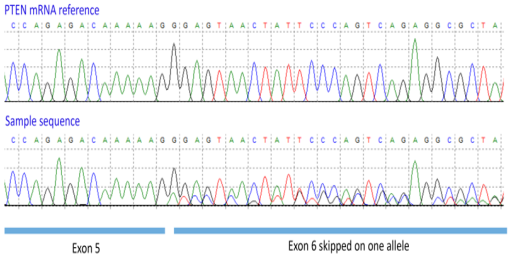
B Exon 4 skipping



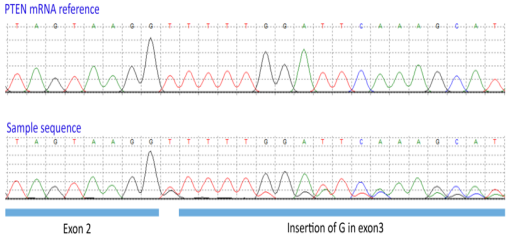
C Exon 5 skipping



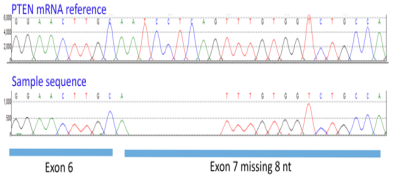
D Exon 6 skipping



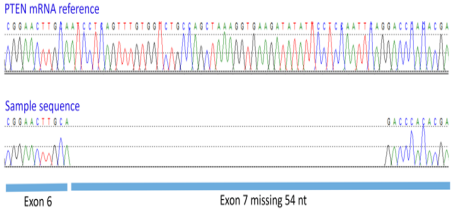
E c.165-2A>G



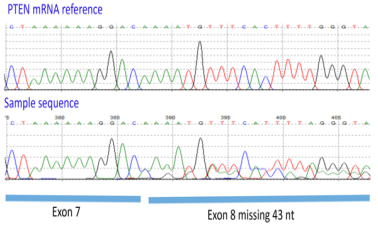
F r.635_642del8



G r.635_688del54



H r.802_844del43

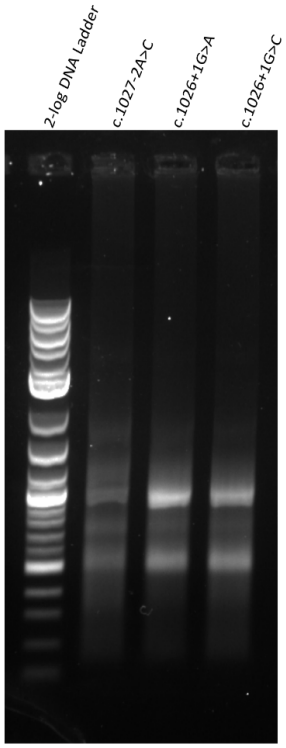


Supplemental Figure 2.

PTEN[2] cDNA sequencing of intron 8 mutations using 3'RACE. **A**: Agarose gel showing the 3'RACE PCR products for the following intron 8 mutations: c.1027-2A>C, c.1026+1G>A, c.1026+1G>C. **B**: Intron 8 mutation c.1027-2A>C results in the mutant transcript premature termination within exon 8 at r.962. Poly (A) is clearly shown. **C**: Intron 8 mutation c. 1026+1G>A results in the mutant transcript with intron 8 retention, extending directly from exon 8 sequence. **D**: Intron 8 mutation c.1026+1G>C results in the mutant transcript with intron 8 retention, extending directly from exon 8 sequence. **E**: RT-PCR using 2 sets of primers with forward primers reside in *PTEN* exon 7 and reverse primers reside in intron 8 retention region as detected in C and D. Agarose gel showing the PCR results from control blood cDNA, HeLa cell cDNA and normal control LBL cDNA. Pooled human genomic DNA (Promega) was used as negative control for genomic DNA contamination of RNA/RT-PCR. Nucleotide numbering uses +1 as the A of ATG translation initiation codon in the reference sequence, with the initiation codon as codon 1.

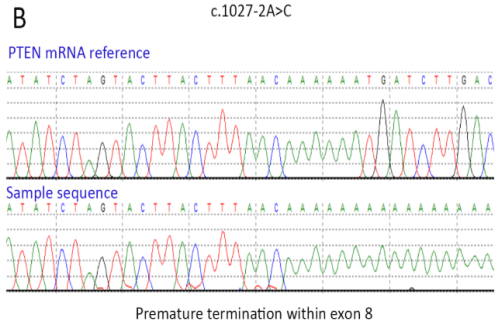
[2] The *PTEN* reference sequence is NC_000010.11

A

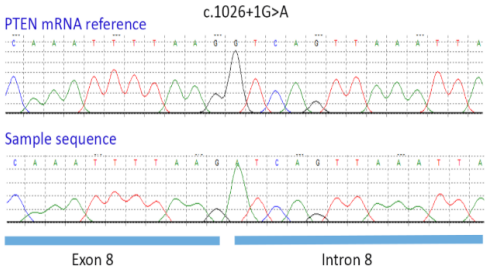

2-log DNA Ladder
c.1027-2A>C
c.1026+1G>A
c.1026+1G>C

3'RACE

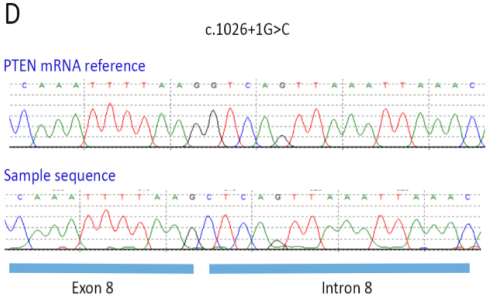
B


c.1027-2A>C
PTEN mRNA reference
Sample sequence
Premature termination within exon 8

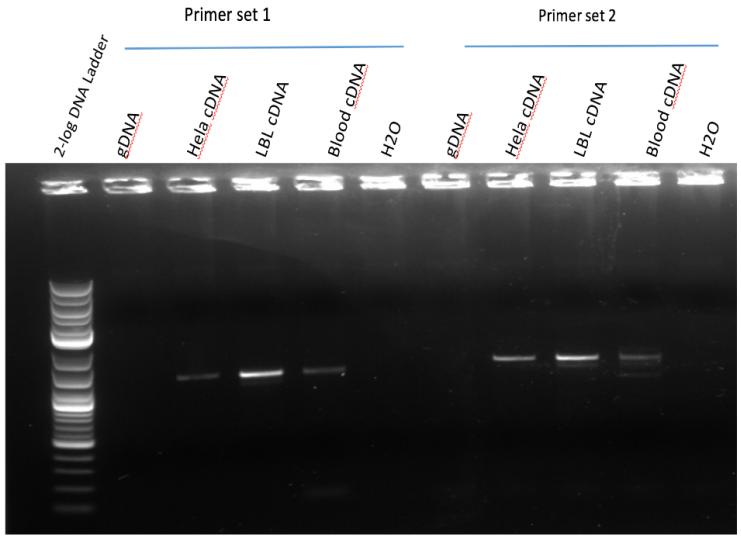
C


c.1026+1G>A
PTEN mRNA reference
Sample sequence
Exon 8
Intron 8

D


c.1026+1G>C
PTEN mRNA reference
Sample sequence
Exon 8
Intron 8

E


Primer set 1
Primer set 2
2-log DNA Ladder
gDNA
Hela cDNA
LBL cDNA
Blood cDNA
H2O
gDNA
Hela cDNA
LBL cDNA
Blood cDNA
H2O

Supplemental Table 1.
Germline *PTEN*[3] intronic variants and clinical features from 61 probands and family members included in RNA processing and protein expression analysis

Patient	Age of Onset	Variation Name	Intron	CC Score	Macrocephaly	Breast cancer	Thyroid cancer	Endometrial carcinoma	Lhermitte-Duclos disease	Other thyroid lesions	Intellectual disability	Hamartomatous intestinal polyps	Fibrocystic disease of the breast	Lipomas	Fibromas	Genitourinary tumors	Genitourinary malformation	Trichilemmoma	Acral keratoses	Papillomatous lesions	Mucosal lesions	Negative for phenotypes of interest
4	48	c.165-2A>G	2	8										X		X						
4a	52	c.165-2A>G	2	10					X													
5	59	c.165-1G>C	2	NA	X					X		X		X								
6	64	c.210-1G>A	3	11	X					X						X						
7	59	c.210-1G>A	3	18	X	X				X			X			X						
7a	53	c.210-1G>A	3	14	X		X							X				X				
7b	19	c.210-1G>A	3	26	X		X			X												
19	63	c.209+4_209+7delAGTA	3	NA	X	X			X	X										X	X	
20	11	c.209+5G>A	3	18	X					X				X								
21	NA	c.210-7_210-3delCTTTT	3	NA				X		X												
23	55	c.253+1G>A	4	23						X		X	X	X		X		X		X	X	
24	49	c.253+1G>T	4	29	X	X				X		X				X						
25	4	c.253+5G>T	4	8	X									X								

[3] The *PTEN* reference sequence is NC_000010.11

26	34	c.253+5G>A	4	15	X					X						X						
30	54	c.493-2A>G	5	19	X	X		X		X						X			X		X	
31	7	c.492+1G>T	5	7	X																	
3	16	c.80-3C>G c.634+4A>T	1&6	21	X					X											X	
34	34	c.634+2T>C	6	37	X				X			X		X	X			X		X	X	
34a	10	c.634+2T>C	6	17	X							X								X		
34b	NA	c.634+2T>C	6	7	X									X								
35	25	c.635-1G>C	6	27	X					X		X		X						X		
36	45	c.634+1G>C	6	27	X	X							X	X		X					X	
37	45	c.634+5G>C	6	68	X		X		X	X		X		X	X				X		X	
38	55	c.802-2A>T	7	44	X							X		X		X		X			X	
39	27	c.802-2A>G	7	38	X					X		X		X						X	X	
46	35	c.1026+1G>C	8	37	X				X	X		X		X								
47	54	c.1026+1G>C	8	33	X					X		X		X		X			X		X	
48	21	c.1026+1G>A	8	37	X					X							X				X	
48a	45	c.1026+1G>A	8	20	X	X				X				X		X					X	
48b	66	c.1026+1G>A	8	27	X	X	X			X			X								X	
49	43	c.1027-2A>C	8	25	X			X		X			X	X		X		X			X	
1	50	c.79+7A>G	1	11		X	X									X						
2	60	c.79+7A>G	1	2									X			X						
3a	15	c.80-3C>G	1	NA																		X
3b	46	c.80-3C>G	1	NA	X																	
8	60	c.210-7_210-3delCTTTT	3	15	X	X				X				X								
9	29	c.210-7_210-3delCTTTT	3	14										X		X					X	
9a	34	c.210-7_210-3delCTTTT	3	7																		X
9b	69	c.210-7_210-3delCTTTT	3	NA																		X
10a	54	c.210-7_210-3delCTTTT	3	NA		X							X									

11	47	c.210-7_210-3delCTTTT	3	0																		X
12	NA	c.210-7_210-3delCTTTT	3	NA	X																	
13	59	c.210-7_210-3delCTTTT	3	10	X					X												
14	55	c.210-7_210-3delCTTTT	3	7	X	X							X									
10	50	c.210-7_210-3delCTTTT	3	17	X		X							X		X						
15	49	c.210-4_210-1delTTAG	3	42	X	X				X		X				X			X	X	X	
16	71	c.210-9T>C	3	NA		X																
17	50	c.210-39A>G	3	3		X							X	X	X	X						
18	40	c.210-39A>G	3	6			X						X			X						
22	NA	c.210-8delT	3	NA		X																
27	36	c.254-51A>T,c.254-72A>T	4	15	X	X							X									
28	59	c.254-38dupT	4	2			X							X			X					
29	37	c.254-38dupT	4	NA		X	X			X												
32	8	c.493-31A>G	5	17								X										
33	47	c.493-52A>G	5	NA		X																
40	56	c.802-51_802-14del38	7	NA		X																
41	54	c.802-51_802-14del38	7	NA		X										X						
42	58	c.802-51_802-14del38	7	NA		X																
43	48	c.802-51_802-14del38	7	0								X										
44	50	c.802-51_802-14del38	7	2		X																
45	38	c.802-51_802-14del38	7	6						X			X	X								

Nucleotide numbering uses +1 as the A of ATG translation initiation codon in the reference sequence, with the initiation codon as codon 1. Probands are denoted by a unique patient number. Family members are denoted by a combination of the proband's unique patient number and a lowercase letter.

Supplemental Table 2. List of primers used in the study

Name	Sequence
969F(5'UTR)	CATTTCCATCCTGCAGAAGAAG
2777R	GGTCCAGAGTCCAGCATAAA
1438F(E5)AS	GGGACGAACTGGTGTAATGATATG
2281(3'UTR)	AGGTCCATTTTCAGTTTATTCAAGTTTA
969F(5'UTR)	CATTTCCATCCTGCAGAAGAAG
1402-1422R(E5)	TCGTCCCTTTCCAGCTTTAC
1736F(E7)	GGACCCACACGACGGGAA
2309-2281R(3'UTR)	AGGTCCATTTTCAGTTTATTCAAGTTTA
E7F-(1704)	GCCAGCTAAAGGTGAAGATA
IVS8R+50	CACACATCACATACATACAAGTC
E7F-1752	CGACGGGAAGACAAGTTCATGTAC
IVS8R+183	TGACGCTGTGTACATTGGGTATTT
PTEN 3'RACE-GSP	GATTACGCCAAGCTTGGACCCACACGACGGGAAGACAAGTTC

Supplemental Table 3.
Germline *PTEN*[4] exonic mutations and clinical features from 24 probands and family members.

Patient	Age of Onset	Variation Name	Exon	CC Score	Macrocephaly	Breast cancer	Thyroid cancer	Endometrial carcinoma	Lhermitte-Duclos disease	Other thyroid lesions	Intellectual disability	Hamartomatous intestinal polyps	Fibrocystic disease of the breast	Lipomas	Fibromas	Genitourinary tumors	Genitourinary malformation	Trichilemmoma	Acral keratoses	Papillomatous lesions	Mucosal lesions	Negative for phenotypes of interest
101	35	c.264T>A (p.Tyr88*)	5	35	X	X				X		X	X									
102	40	c.264T>G (p.Tyr88*)	5	51	X		X		X	X		X				X					X	
103	62	c.386G>A (p.Gly129Glu)	5	37	X					X			X	X								
104	42	c.388C>G (p.Arg130Gly)	5	16	X	X				X			X	X		X			X	X		
105	61	c.388C>T (p.Arg130*)	5	38	X			X		X		X			X	X	X	X		X	X	
106	57	c.388C>T (p.Arg130*)	5	13	X	X	X															
107	72	c.389G>A (p.Arg130Gln)	5	18	X	X				X						X		X		X		
108	46	c.389G>A (p.Arg130Gln)	5	43	X	X				X		X	X			X			X			
109	39	c.389G>A (p.Arg130Gln)	5	37	X	X				X			X			X		X	X	X	X	
110	68	c.406T>C (p.Cys136Arg)	5	69	X			X		X		X	X	X	X			X	X	X	X	
111	52	c.406T>C (p.Cys136Arg)	5	18		X	X	X		X			X	X								
112	46	c.406T>C (p.Cys136Arg)	5	18	X	X				X			X								X	
113	37	c.406T>C (p.Cys136Arg)	5	20	X					X												
114	NA	c.407_423del17 (p.Cys136Serfs*38)	5	14		X						X										
115	43	c.407G>A (p.Cys136Tyr)	5	37	X					X		X	X			X				X	X	
116	6	c.420_421insA (p.His141Thrfs*39)	5	13	X																	

[4] The *PTEN* reference sequence is NC_000010.11

117	52	c.422A>G (p.His141Arg)	5	17	X			X								X						
118	39	c.463T>A (p.Tyr155Asn)	5	28	X					X						X					X	
118a	12	c.463T>A (p.Tyr155Asn)	5	19	X																X	
119	30	c.562T>C (p.Tyr188His)	6	10	X																	
119a	4	c.562T>C (p.Tyr188His)	6	7	X																	
120	31	c.609_611delTCCinsATAAAT (p.Pro204*)	6	41	X		X			X		X	X		X							
121	60	c.613A>G (p.Met205Val)	6	12	X	X										X						
122	24	c.632delG (p.Cys211fs)	6	37	X					X	X	X								X	X	

Nucleotide numbering uses +1 as the A of ATG translation initiation codon in the reference sequence, with the initiation codon as codon 1. Probands are denoted by a unique number. Family members are denoted by a combination of the proband's unique patient number and a lowercase letter.