The American Journal of Human Genetics

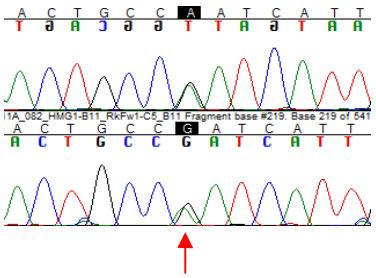
Supplemental Data

A Recurrent Mutation in *PARK2* Is Associated with Familial Lung Cancer

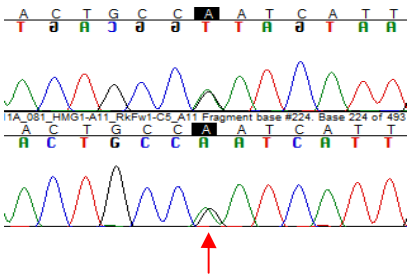
Donghai Xiong, Yian Wang, Elena Kupert, Claire Simpson, Susan M. Pinney, Colette R. Gaba, Diptasri Mandal, Ann G. Schwartz, Ping Yang, Mariza de Andrade, Claudio Pikielny, Jinyoung Byun, Yafang Li, Dwight Stambolian, Margaret R. Spitz, Yanhong Liu, Christopher I. Amos, Joan E. Bailey-Wilson, Marshall Anderson, and Ming You

Family A (c.823C>T [p.Arg275Trp]):

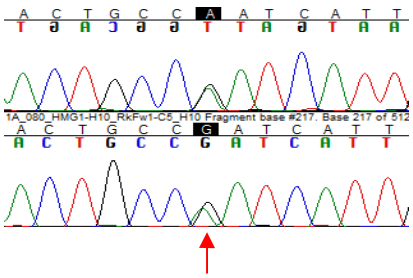
Subject III-1



Subject III-4

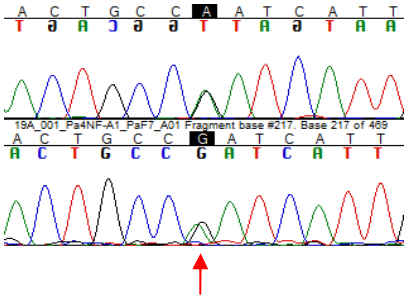


Subject IV-18

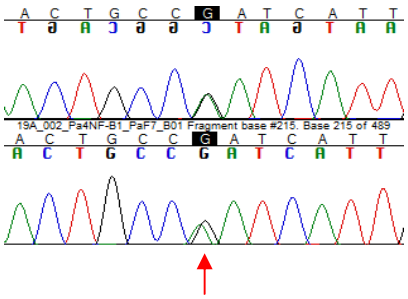


Family B (c.823C>T [p.Arg275Trp]):

Subject III-6

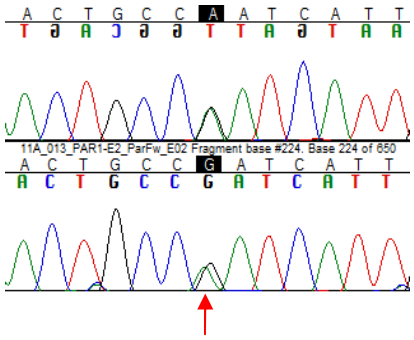


Subject III-7

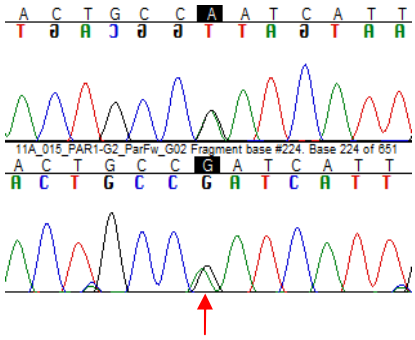


Family C (c.823C>T [p.Arg275Trp]):

Subject II-3



Subject II-1



Family D (c.823C>T [p.Arg275Trp])

Subject II-2

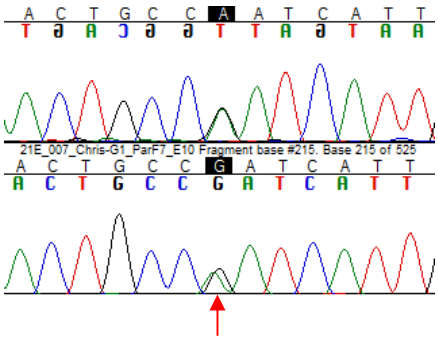


Figure S1. Sanger sequencing chromatograms showing germline *PARK2* c.823C>T (p.Arg275Trp) mutations in familial lung cancer patients. Mutations ~~were~~ indicated by arrows. Both forward and reverse sequence chromatograms ~~were~~ given for each c.823C>T (p.Arg275Trp) mutation identified in the patients.

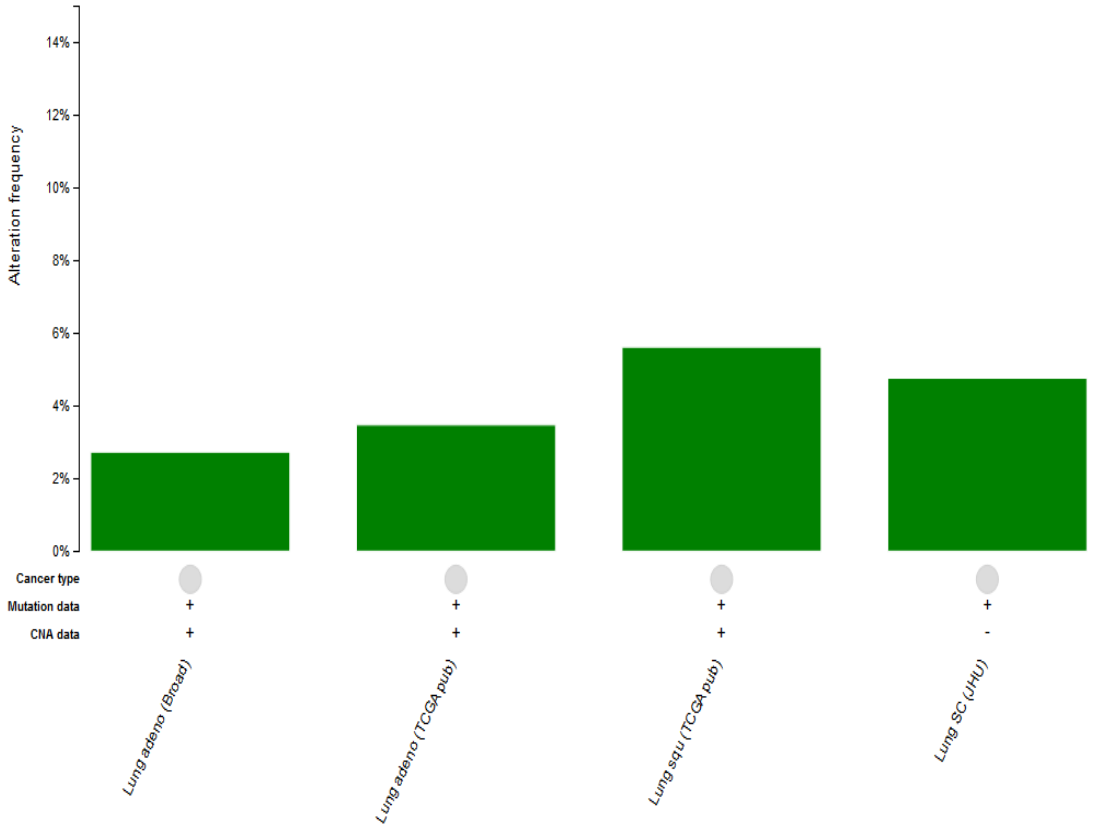
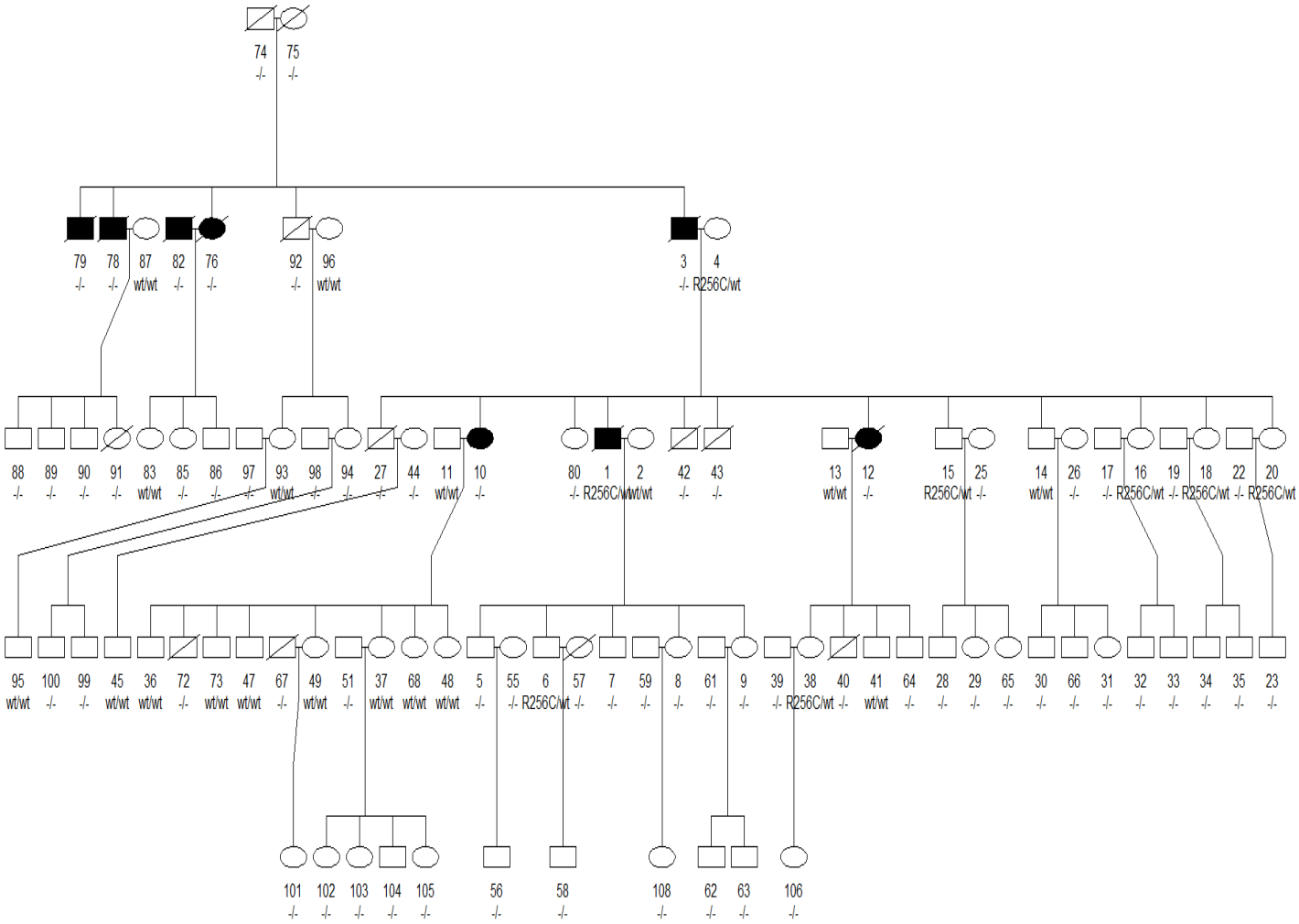
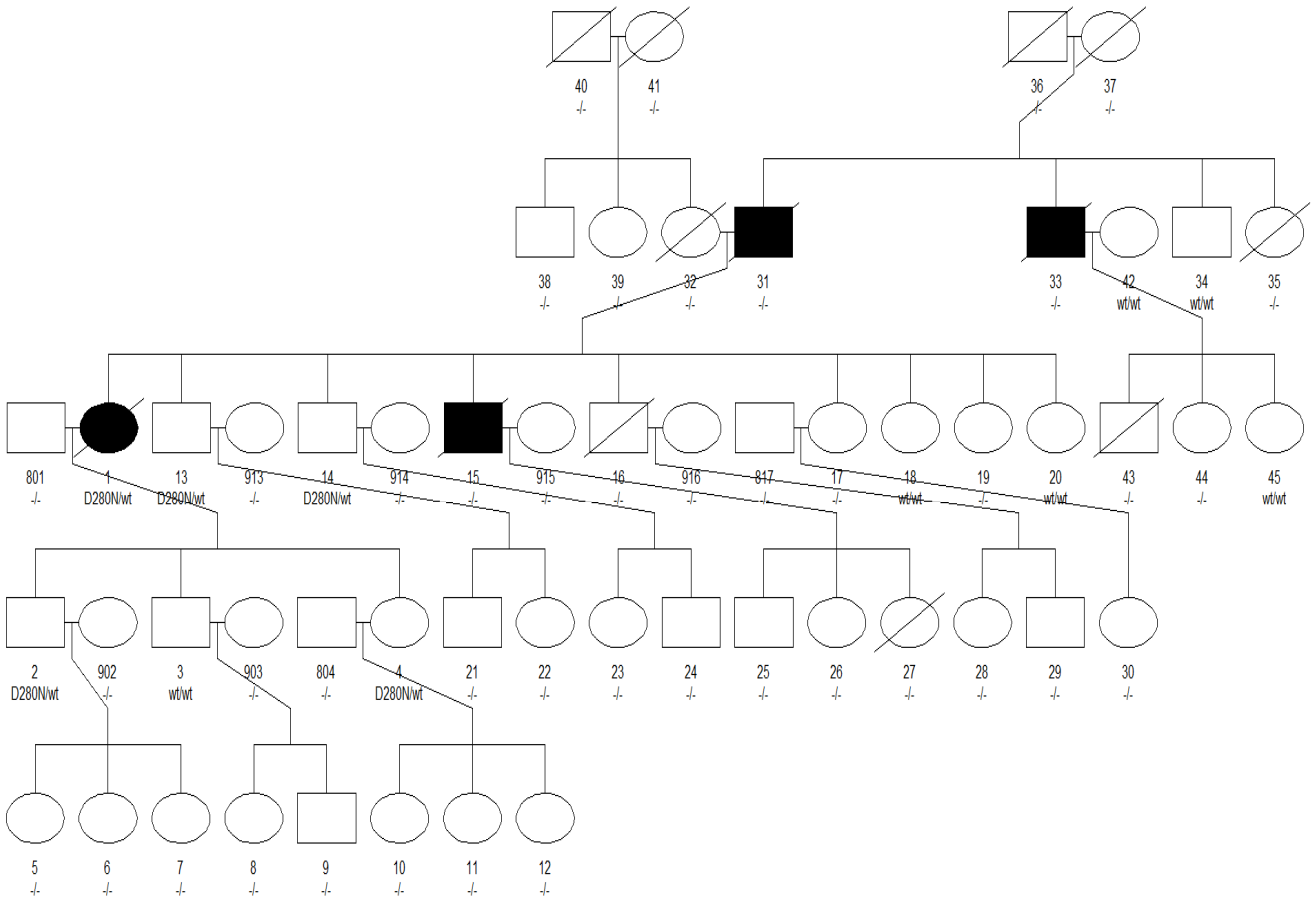


Figure S2. Somatic mutation rates of *PARK2* in different types of lung cancers according to the cBioPortal for Cancer Genomics (http://www.cbioportal.org/public-portal/). Lung adeno: lung adenocarcinoma; Lung squ: lung squamous cell carcinoma; Lung sc: small cell lung cancer.

A) c.766C>T (p.Arg256Cys) mutation (labeled as ‘R256C’) in Family E



B) c.838G>A (p.Asp280Asn) mutation (labeled as 'D280N') in Family F



C) c.668C>T (p.Ser223Leu) mutation (labeled as 'S223L') in Family G

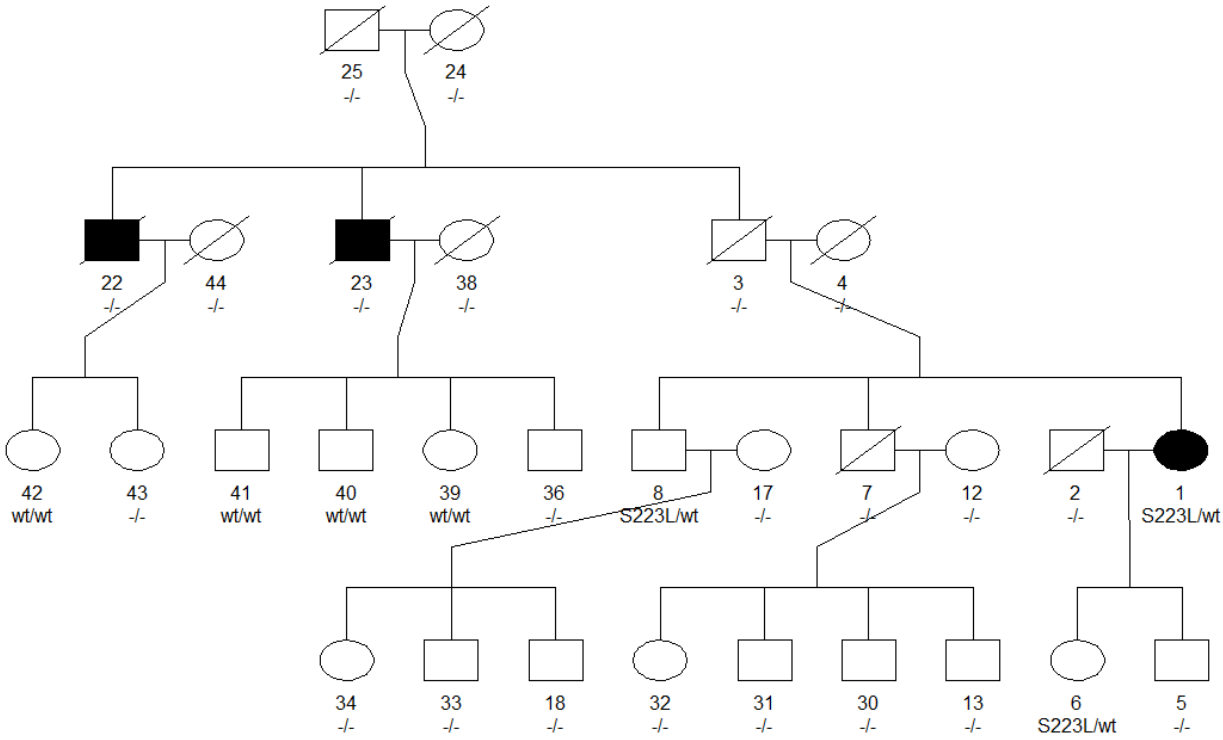


Figure S3. The three pedigrees with corresponding genotype information for the *PARK2* mutations - c.766C>T (p.Arg256Cys), c.838G>A (p.Asp280Asn), c.668C>T (p.Ser223Leu). A) c.766C>T (p.Arg256Cys) mutation in the family E; B) c.838G>A (p.Asp280Asn) mutation in the family F; C) c.668C>T (p.Ser223Leu) mutation in the family G. "-/-" denotes no DNA sample available for the corresponding subject. "wt" denotes wild-type allele.

Table S1. Phenotype information and genotypes of *PARK2* c.823C>T (p.Arg275Trp) for the members in the three multiple-lung-cancer families.

Family A (49 subjects with smoking information)

Subject ID	c.823C>T (p.Arg275Trp)	Lung cancer status (age at diagnosis for cases)	Smoking status
III-1	Mutated	Affected (60)	Former smoker
IV-18	Mutated	Affected (69)	Former smoker
III-10	Mutated	Affected (65)	Former smoker
III-7	Mutated	Affected (62)	Former smoker
III-4	Mutated	Affected (38)	Never smoker
IV-9	Unknown	Affected (39)	Former smoker
IV-7	Unknown	Affected (47)	Former smoker
V-1	Wild type	Affected with SCLC (44)	Current smoker
IV-12	Mutated	Currently unaffected but has leukemia	Current smoker
V-5	Mutated	Currently unaffected	Never smoker
IV-6	Mutated	Currently unaffected	Current smoker
III-5	Mutated	Currently unaffected	Unknown
IV-3	Mutated	Currently unaffected	Current smoker
IV-14	Mutated	Currently unaffected	Current smoker
V-4	Mutated	Currently unaffected	Never smoker
I-1	Unknown	Currently unaffected	Unknown
I-2	Unknown	Currently unaffected	Unknown
II-2	Unknown	Currently unaffected	Unknown
II-1	Unknown	Currently unaffected	Unknown
II-4	Unknown	Currently unaffected	Never smoker
II-3	Unknown	Currently unaffected	Never smoker
II-5	Unknown	Currently unaffected	Unknown
II-6	Unknown	Currently unaffected	Unknown
III-2	Unknown	Currently unaffected	Unknown
III-3	Wild type	Currently unaffected	Never smoker
III-8	Unknown	Currently unaffected	Former smoker
III-6	Wild type	Currently unaffected	Never smoker
III-9	Wild type	Currently unaffected	Never smoker
III-11	Unknown	Currently unaffected	Unknown

Family A (continue)

Subject ID	c.823C>T (p.Arg275Trp)	Lung cancer status (age at diagnosis for cases)	Smoking status
IV-20	Wild type	Currently unaffected	Current smoker
IV-21	Wild type	Currently unaffected	Current smoker
IV-22	Wild type	Currently unaffected	Never smoker
IV-23	Wild type	Currently unaffected	Current smoker
V-2	Wild type	Currently unaffected	Current smoker
V-3	Wild type	Currently unaffected	Current smoker
V-6	Wild type	Currently unaffected	Never smoker

Family B (25 subjects with smoking information)

Subject ID	c.823C>T (p.Arg275Trp)	Lung cancer status (age at diagnosis for cases)	Smoking Status
II-2	Unknown	Affected (NA)	Never smoker
II-4	Unknown	Affected (NA)	Former smoker
III-7	Mutated	Affected (61)	Never smoker
III-4	Unknown	Affected (NA)	Never smoker
III-6	Mutated	Affected (67)	Never smoker
III-8	Wild type	Currently unaffected	Never smoker
IV-9	Mutated	Currently unaffected	Never smoker
IV-8	Mutated	Currently unaffected	Never smoker
II-1	Unknown	Currently unaffected	Never smoker
III-9	Wild type	Currently unaffected	Never smoker
III-2	Wild type	Currently unaffected	Never smoker
I-1	Unknown	Currently unaffected	Former smoker
I-2	Unknown	Currently unaffected	Never smoker
II-3	Unknown	Currently unaffected	Never smoker
III-5	Wild type	Currently unaffected	Former smoker
IV-7	Mutated	Currently unaffected	Never smoker
IV-6	Wild type	Currently unaffected	Never smoker
III-3	Wild type	Currently unaffected	Never smoker
IV-5	Wild type	Currently unaffected	Never smoker

III-12	Unknown	Currently unaffected	Unknown
IV-1	Wild type	Currently unaffected	Current smoker
IV-2	Wild type	Currently unaffected	Current smoker
IV-5	Wild type	Currently unaffected	Never smoker
IV-4	Wild type	Currently unaffected	Never smoker
IV-8	Wild type	Currently unaffected	Former smoker
IV-10	Wild type	Currently unaffected	Current smoker
IV-11	Unknown	Currently unaffected	Never smoker
IV-13	Unknown	Currently unaffected	Current smoker
IV-15	Wild type	Currently unaffected	Never smoker
IV-16	Wild type	Currently unaffected	Never smoker
IV-17	Unknown	Currently unaffected	Unknown
IV-19	Wild type	Currently unaffected	Current smoker

IV-3	Wild type	Currently unaffected	Never smoker
IV-4	Wild type	Currently unaffected	Never smoker
III-1	Unknown	Currently unaffected	Never smoker
IV-2	Wild type	Currently unaffected	Never smoker
IV-1	Wild type	Currently unaffected	Never smoker

Family C (4 subjects with smoking information)

Subject ID	c.823C>T (p.Arg275Trp)	Lung cancer status (age at diagnosis for cases)	Smoking Status
I-1	Wild type	Affected and also has prostate cancer (56)	Never smoker
II-1	Mutated	Affected (57)	Never smoker
II-2	Wild type	Affected (63)	Former smoker
II-3	Mutated	Affected (59)	Unknown

Table S2. Characteristics of the eight lung cancer patients in Family A.

Individual ID	Pathology	Age at onset	Smoking status
III-1	Lung squamous cell carcinoma	60	Former smoker
III-7	Lung squamous cell carcinoma	62	Former smoker
III-10	Lung cancer (type unknown NSCLC)	65	Former smoker
III-4	Lung carcinoid	38	Never smoker
IV-18	Lung adenocarcinoma	69	Former smoker
IV-9	Small cell lung cancer	39	Former smoker
IV-7	Small cell lung cancer	47	Former smoker
V-1	Small cell lung cancer	44	Current smoker

Table S3. The list of somatic mutations of *PARK2* identified by the sequencing projects of lung tumors from TCGA, Broad Institute and John Hopkins' studies.

Case ID	Cancer Study	AA Change	Type
LUAD-CHTN-MAD06-00490	Lung adenoCA (Broad)	H373D	Missense
LUAD-GU4I3	Lung adenoCA (Broad)	G336F	Missense
LUAD-B00915	Lung adenoCA (Broad)	C449F	Missense
LUAD-S01467	Lung adenoCA (Broad)	R156*	Nonsense
LUAD-B02594	Lung adenoCA (Broad)	G179_splice	Splice
TCGA-78-7155	Lung adenoCA (TCGA pub)	E322*	Nonsense
TCGA-05-4410	Lung adenoCA (TCGA pub)	G354R	Missense
TCGA-78-7159	Lung adenoCA (TCGA pub)	A138P	Missense
TCGA-64-5775	Lung adenoCA (TCGA pub)	P343Q	Missense
TCGA-44-7670	**Lung adenoCA (TCGA pub)**	**R275P**	**Missense**
TCGA-55-7907	Lung adenoCA (TCGA pub)	E353*	Nonsense
TCGA-05-4410	Lung adenoCA (TCGA pub)	G213V	Missense
TCGA-49-4501	Lung adenoCA (TCGA pub)	G429_splice	Splice
TCGA-05-4390	Lung adenoCA (TCGA pub)	W453L	Missense
TCGA-66-2754	Lung squCA (TCGA pub)	E300G	Missense
TCGA-60-2713	Lung squCA (TCGA pub)	I229F	Missense
TCGA-18-3416	Lung squCA (TCGA pub)	R97L	Missense
TCGA-66-2757	Lung squCA (TCGA pub)	Q347H	Missense
TCGA-66-2763	Lung squCA (TCGA pub)	A405P	Missense
TCGA-66-2777	Lung squCA (TCGA pub)	A138_splice	Splice
TCGA-34-2608	Lung squCA (TCGA pub)	V186I	Missense
TCGA-21-1081	Lung squCA (TCGA pub)	H200P	Missense
TCGA-21-5787	Lung squCA (TCGA pub)	K211N	Missense
TCGA-18-3409	Lung squCA (TCGA pub)	P113S	Missense
2334201	Lung small cell CA (JHU)	T222A	Missense
134430	Lung small cell CA (JHU)	C95S	Missense

Table S4. Primers for PCR and Sanger sequencing of all the 12 exons of the *PARK2* gene. Primers were designed using the Primer3Plus program (http://primer3plus.com). *Genomic coordinates were based on human assembly hg19.

Forward Primer	Reverse Primer	*Amplicon Chr Location	Amplicon coverage
5'-GGTCTTCATGAGAACGCTCAG-3'	5'-CGTTAGAACTACGACTCCCAGC-3'	chr6: 163,148,451-163,148,887	whole exon 1
5'-GCCTAGCTCGTGTGTTCTTTCT-3'	5'-TTGCTATCACCATTTAAGGGCT-3'	chr6: 162,864,094-162,864,574	whole exon 2
5'-AACTAAATATGCACCCGGTGAG-3'	5'-CTCGCATTTCATGTTTGACATT-3'	chr6: 162,683,464-162,683,860	whole exon 3
5'-TTTCTTTTCAAAGACGGGTGAT-3'	5'-GATACACTTGCCCGATTCTCTT-3'	chr6: 162,622,058-162,622,563	whole exon 4
5'-GCAATGAGTTTAAATTGGCACA-3'	5'-TGCAGACCACACTTTGAAAATC-3'	chr6: 162,474,866-162,475,341	whole exon 5
5'-AAAATAAAGCAGACACTCCCCA-3'	5'-AGACTCAGGGCCCTTCTAAAAC-3'	chr6: 162,394,180-162,394,658	whole exon 6
5'-CCAGTTCAACACAATTCCTTCA-3'	5'-ACAACCCTCCAGGATTACAGAA-3'	chr6: 162,206,609-162,207,034	whole exon 7
5'-CACACATATCTCCAGCATGGTT-3'	5'-CCAGATCATAATCAGAGGGGAG-3'	chr6: 161,990,241-161,990,715	whole exon 8
5'-TAAAAGTCTGGCCTAGTGGCTC-3'	5'-ACTGGTTAAGCAAGAAATCCCA-3'	chr6: 161,969,606-161,970,133	whole exon 9
5'-GAATGGAACTCTCCATGACCTC-3'	5'-ATCTTGAGGGAAGGAAATGTGA-3'	chr6: 161,807,735-161,808,156	whole exon 10
5'-CTCCTTTAATCCTGGAATCCCT-3'	5'-CCGCCTAGTAGCTGTCTAGCAT-3'	chr6: 161,781,034-161,781,403	whole exon 11
5'-ACAAACTGAAAGGGATTCAGGA-3'	5'-GAGCTGCCCTATTGTGCTTTAT-3'	chr6: 161,770,887-161,771,422	whole exon 12

Table S5. Four additional missense *PARK2* mutations identified in familial lung cancer cases not carrying the c.823C>T (p.Arg275Trp).

Family	Cases in the family	Genomic Position[a]	Genomic mutation	Exon	Protein alteration	Predicted effect of missense mutation			GERP++[e]	dbSNP137	ESP MAF[f]	1000 Genomes MAF[g]
						PolyPhen-2[b]	SIFT[c]	FATHMM[d]				
E	1	Chr. 6: 162206909	c.766C>T	Exon 7	p.Arg256Cys	Probably damaging (1.00)	Deleterious (0.02)	Potentially associated with cancer (-3.28)	5.75	rs150562946	0.0004	0.0005
F	1	Chr. 6: 162206837	c.838G>A	Exon 7	p.Asp280Asn	Probably damaging (1.00)	Tolerated (0.09)	Potentially associated with cancer (-3.23)	5.75	rs72480422	NA	NA
G	1	Chr. 6: 162394400	c.668C>T	Exon 6	p.Ser223Leu	Probably damaging (0.98)	Deleterious (0.01)	Potentially associated with cancer (-3.31)	5.01	NA	NA	NA
H	1	Chr. 6: 161781201	c.1204C>T	Exon 11	p.Arg402Cys	Probably damaging (1.00)	Deleterious (0.00)	Potentially associated with cancer (-3.29)	1.8	rs55830907	0.0018	0.0037

[a]Genomic positions are given according to the hg19 reference assembly; [b]PolyPhen-2 scores 0.85-1 are interpreted as probably damaging, scores 0.2–0.85 as possibly damaging and scores 0-0.2 as benign; [c]SIFT scores range from 0 to 1. The amino acid substitution is predicted damaging if the score is <= 0.05, and tolerated if the score is > 0.05. [d]Predictions with FATHMM scores less than -0.75 indicate that the mutation is potentially associated with cancer, otherwise not associated with cancer; [e]An indication of evolutionary conservation is made if a given site shows GERP++ score > 2; [f]Minor allele frequencies (MAF) are according to the National Heart, Lung, and Blood Institute (NHLBI) GO Exome Sequencing Project (ESP6500SI-V2 release) Exome Variant Server v.0.0.21 (August 2013). [g]Minor allele frequencies (MAF) are according to the 1000 Genomes Project 2012 Aprial data sets.