Supporting Information

Poulogiannis et al. 10.1073/pnas.1009941107

SI Materials and Methods

Bioinformatic Analysis of Genome and Transcriptome Data. The aCGH package in R was used to identify significant DNA copy number (DCN) changes in our collection of 100 sporadic CRCs (1) (Gene Expression Omnibus, accession no. GSE12520). The aCGH analysis of cell lines and liver metastases was derived from published data (2, 3). Chromosome 6 tiling-path array-CGH was used to identify the smallest and most frequently altered regions of DNA copy number change on chromosome 6. An integrative approach was used to correlate expression profiles with genomic copy number data from a SNP array from the same tumors (n = 48) (4) (GSE16125), using Pearson's correlation coefficient analysis to identify the relationships between DNA copy number changes and gene expression of those genes located within the small frequently altered regions of DCN change identified by tiling-path array-CGH on chromosome 6. Finally, the differentially expressed genes between high- and low-*PARK2* expressing tumors, from a previously published dataset (GSE12945) of microdissected CRCs (5), were computed using a linear model, based on an empirical Bayes (eBayes) method (6) embedded within the Limma package. The *P*-values were adjusted for multiple testing (7).

Loss of Heterozygosity (LOH) Analysis of PARK2. Seven microsatellite markers (D6S1550, D6S253, D6S305, D6S955, D6S980, D6S1599, and D6S396) were amplified for LOH analysis within the PARK2 locus using primers that were previously described (8).

MSP of the PARK2 Promoter. CpG sites within the *PARK2* promoter region were detected using the Methprimer software (http://www.urogene.org/methprimer/index.html). Methylation-specific and control primers were designed using the Primo MSP software (http://www.changbioscience.com/primo/primom.html); bisulfite modification of genomic DNA was performed as described previously (9). All tumor DNA samples from primary CRC tumors (n = 100) and CRC lines (n = 5), as well as those from the leukemia cell lines KG-1a (acute myeloid leukemia, AML), U937 (acute lymphoblastic leukemia, ALL), and Raji (Burkitt lymphoma, BL) were screened as part of this analysis. *SssI* DNA, composed of mixed unmethylated and methylated human sperm DNA treated with *SssI* methyltransferase, was used as a positive control (Qiagen).

Immunohistochemistry. Immunostaining was performed on 4-μm sections with the rabbit Vectastain Elite ABC horseradish peroxidase kit (Vector Laboratories) using primary antibodies: β-catenin (1:50; Cell Signaling Technology, no. 9587) and Parkin (1:75; Epitomics, no. 1679–1).

1. Poulogiannis G, et al. (2010) Prognostic relevance of DNA copy number changes in colorectal cancer. *J Pathol* 220:338–347.
2. Douglas EJ, et al. (2004) Array comparative genomic hybridization analysis of colorectal cancer cell lines and primary carcinomas. *Cancer Res* 64:4817–4825.
3. Mehta KR, et al. (2005) Fractional genomic alteration detected by array-based comparative genomic hybridization independently predicts survival after hepatic resection for metastatic colorectal cancer. *Clin Cancer Res* 11:1791–1797.
4. Reid JF, et al. (2009) Integrative approach for prioritizing cancer genes in sporadic colon cancer. *Genes Chromosomes Cancer* 48:953–962.
5. Staub E, et al. (2009) An expression module of WIPF1-coexpressed genes identifies patients with favorable prognosis in three tumor types. *J Mol Med* 87:633–644.
6. 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.
7. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Ser B* 57:289–300.
8. Denison SR, Callahan G, Becker NA, Phillips LA, Smith DI (2003) Characterization of FRA6E and its potential role in autosomal recessive juvenile parkinsonism and ovarian cancer. *Genes Chromosomes Cancer* 38:40–52.
9. Herman JG, Graff JR, Myöhänen S, Nelkin BD, Baylin SB (1996) Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc Natl Acad Sci USA* 93:9821–9826.
10. Mata IF, Lockhart PJ, Farrer MJ (2004) Parkin genetics: One model for Parkinson's disease. *Hum Mol Genet* 13 (Spec No 1):R127–R133.

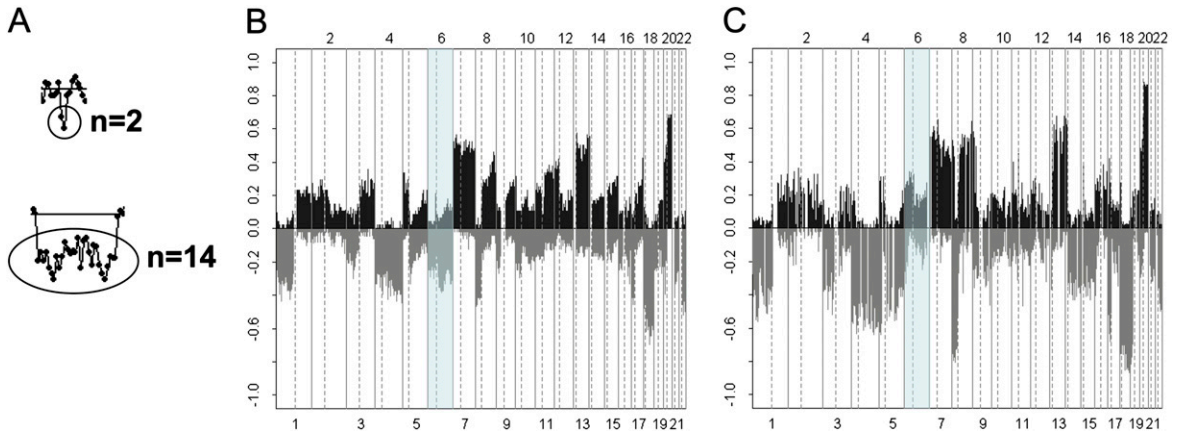

Fig. S1. (*A*) Examples of 2 (*Upper*) versus 14 (*Lower*) consecutive clones that are deleted in CRC samples as detected by chromosome 6 tiling-path array-CGH. (*B* and *C*) Overall frequencies of DNA copy number alterations (*y* axis) plotted by chromosome number (*x* axis) as assessed by 1-Mb array-CGH in 49 CRC lines (2) (*B*) and 51 hepatic resections of metastatic CRC (3) (*C*), respectively.

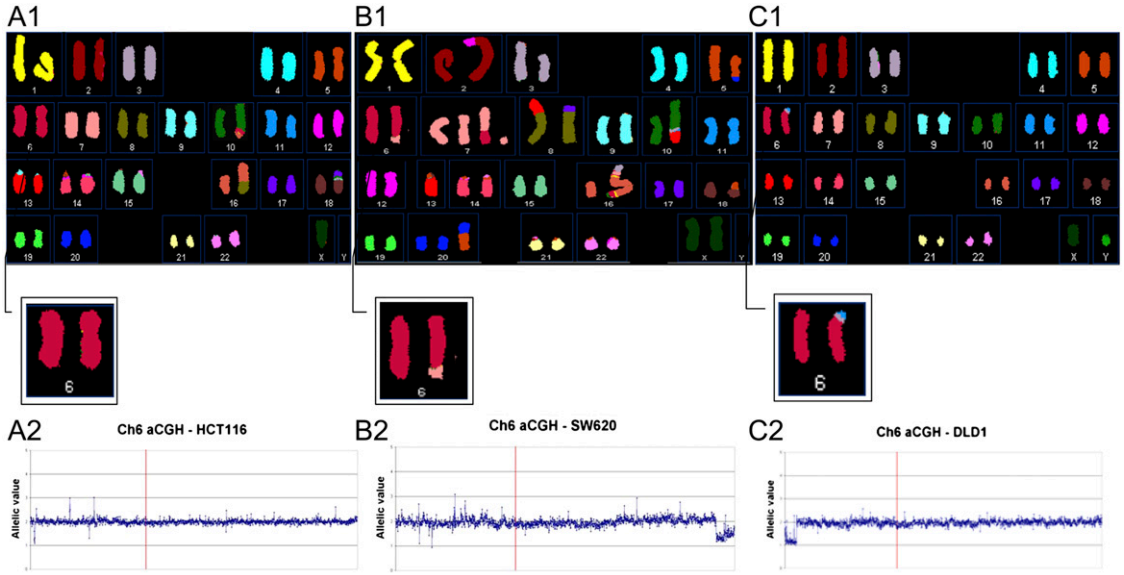

Fig. S2. Examples of spectral karyotypes of CRC cell lines produced by SKY analysis and their relation with the chromosome 6 tiling-path array-CGH data. (*A1–C1*) SKY images of complete metaphases: (*A*) HCT116, (*B*) SW620, and (*C*) DLD1. (*A2–C2*) Copy number plots of the corresponding chromosome 6 tiling-path array-CGH profiles for HCT116, SW620, and DLD1.

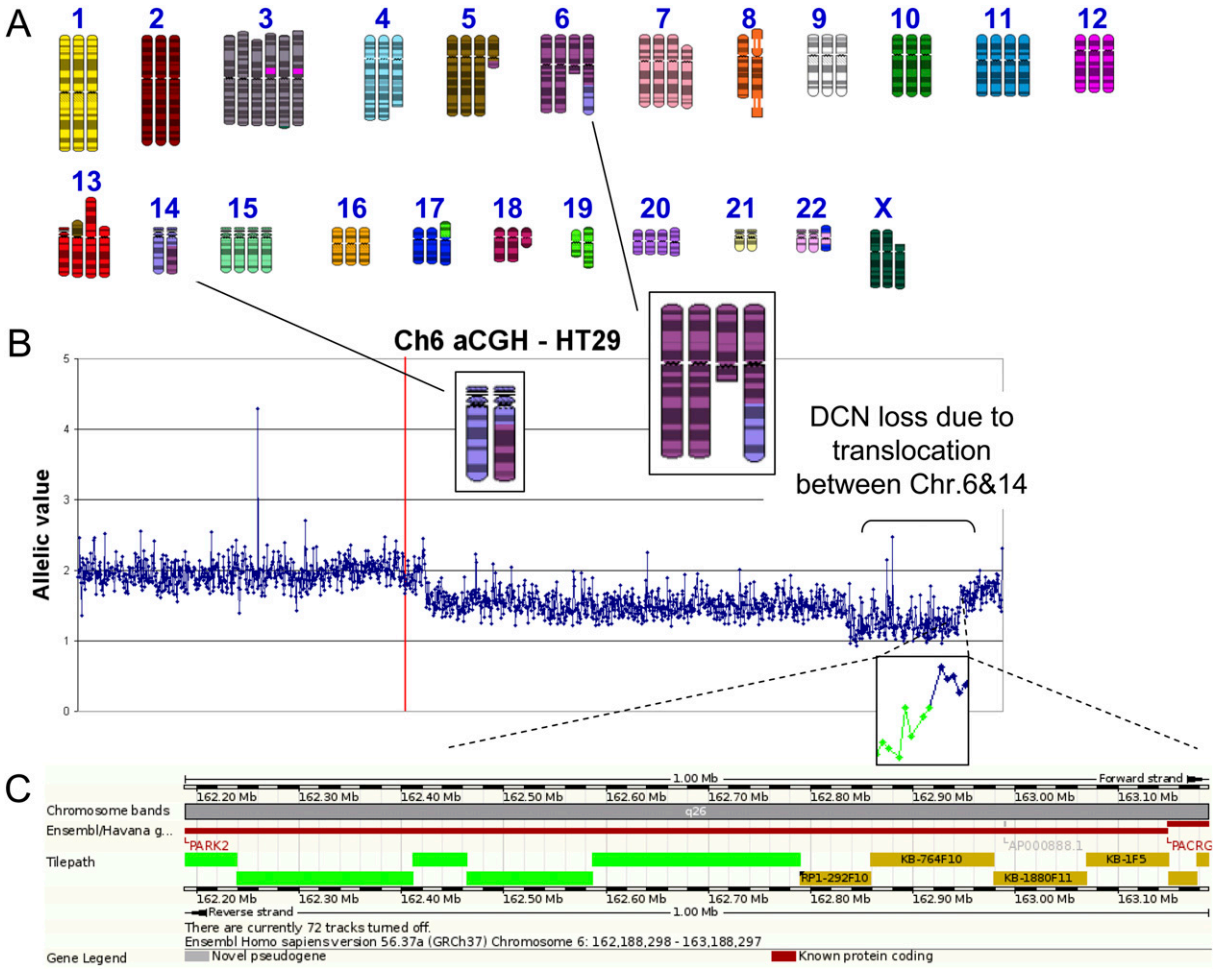

Fig. S3. DCN changes involving the distal part of 6q in CRC cell line HT29. (*A*) SKYGRAM indicating a complex translocation between chromosomes 6 and 14. (*B*) Linear scale plot of DNA copy number changes assessed by chromosome 6 tiling-path aCGH, indicating a deletion due to translocation between chromosomes 6 and 14. (*C*) Genomic location of the region mapping to the distal 6q breakpoint that lies within the *PARK2* gene.

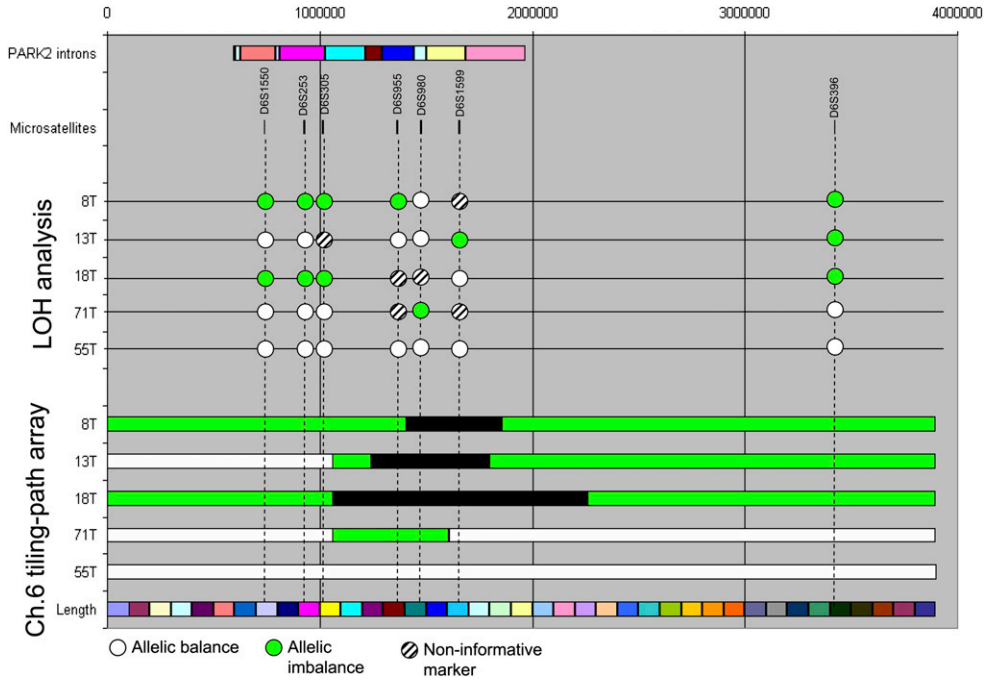

Fig. S4. Comparison of allelic loss data in five CRC primary tumors (including the three with homozygous deletion of *PARK2*) as determined by LOH analysis of seven microsatellite markers and tiling-path array-CGH analysis of chromosome 6 across the commonly deleted region. Each horizontal line represents a single case (identified on the *y* axis) (using the same aCGH color coding system described in Fig. 2). The minimal region of loss within the *PARK2* gene (with exons and introns shown schematically at the top) is indicated by both techniques as allelic balance identified by LOH can be consistent with either diploid biallelic retention or homozygous deletion.

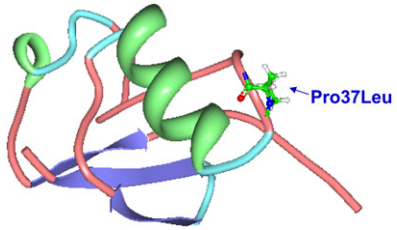

Fig. S5. Ribbon diagram of the PARKIN protein UBL domain, indicating the location of the Pro37Leu (211C > T) somatically mutated amino acid substitution, identified both in the LoVo CRC line and some AR-JP Parkinson disease patients (10).

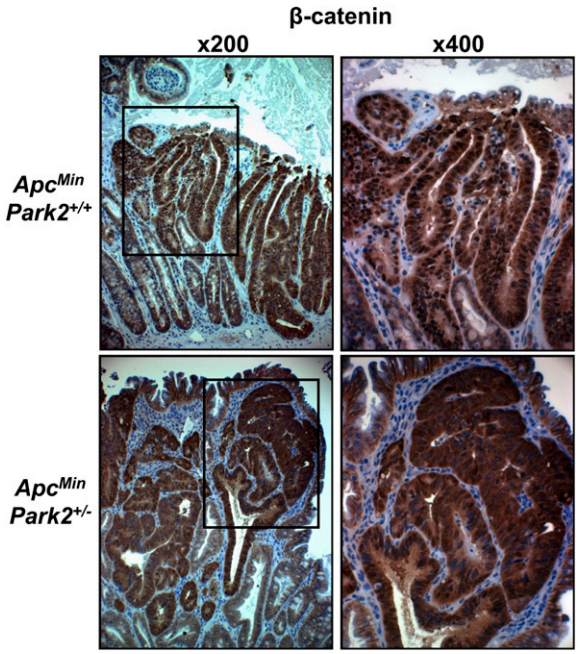

Fig. S6. Representative photomicrographs of β-catenin immunohistochemistry in 4 μm-thick sections containing intestinal adenomas from *Apc*$^{+/Min}$/*Park2*$^{+/+}$ and *Apc*$^{+/Min}$/*Park2*$^{+/-}$ mice, showing greater β-catenin accumulation in adenomas relative to adjacent normal epithelium with nuclear localization. Original magnification 200× (*Upper Left* and *Lower Left*) and 400× (*Upper Right* and *Lower Right*).

Table S1. Correlation statistics between copy number and expression level of the genes located within the small regions of copy number change identified in chromosome 6

Region	Clone name	Band	Change	Candidate genes	Correlation coefficient	*P* value
1	RP3-416J7	6p25.3	G/L	*No genes*	—	—
	RP11-328C17			*DUSP22*	0.5143	0.0013
	RP11-233K4			*IRF4*	0.3765	0.0236
	RP11-164H16			*EXOC2*	0.3371	0.044
	RP1-20B11					
	RP11-532F6			*HUS1B*	NA	NA
2	RP1-118B18	6p25.3	L	*FOXC1*	0.1537	0.3709
	RP11-265E5			*GMDS*	0.0994	0.564
	RP1-279I9					
	RP11-82M9					
	RP11-612M16					
	RP1-206F19					
	RP1-33B19					
	RP11-420G6	6p25.2	G	*MYLK4*	0.5148	0.0013
				RP11-420G6.1	NA	NA
				WRNIP1	0.1425	0.4071
				WRNIP1	0.1328	0.4401
				SERPINB1	0.4133	0.0122
	RP1-90J20			*SERPINB9*	0.4083	0.0134
				SERPINB6	0.0748	0.6646
				NQO2	0.3433	0.0403
	RP1-40E16			*RIPK1*	0.1561	0.3633
				BPHL	NA	NA
				TUBB2A	0.0552	0.749
	RP11-506K6		G/L	*TUBB2B*	0.0484	0.9062
				PSMG4	NA	NA
	RP11-15N12			*SLC22A23*	0.0484	0.779
3	RP11-716O23	6p24.2–24.1	L	*TMEM170B*	NA	NA
	RP11-679B17					
4	XXbac-BPG8G10	6p22.1	L	*BAT1P1*	NA	NA
	XXbac-BPG249D20	6p21.33	L	*MCCD1P1*	NA	NA
				GNL1	NA	NA
				PRR3	NA	NA
				ABCF1	NA	NA
				PPP1R10	NA	NA
				MRPS1B	NA	NA
				C6orf134	NA	NA
	XXbac-BPG27H4			*DHX16*	NA	NA
	XXbac-BPG118E17		G	*VARS2*	NA	NA
				GTF2H4	−0.2117	0.2151
				SFTA2	NA	NA
				DPCR1	NA	NA
				C6orf205	NA	NA
	XXbac-BPG248L24		G/L	*HLA-C*	0.3913	0.0183
				HLA-B	NA	NA
	XXbac-BPG181B23			*MICA*	NA	NA

Table S1. Cont.

Region	Clone name	Band	Change	Candidate genes	Correlation coefficient	*P* value
5	XXbac-BPG296P20	6p21.33	L	*ATP6V1G2*	NA	NA
				MCCD1	NA	NA
				BAT1	NA	NA
				NFKBIL1	NA	NA
				LTA	NA	NA
				TNF	NA	NA
				LTB	NA	NA
				LST1	NA	NA
				NCR3	NA	NA
				BAT2	NA	NA
				BAT3	NA	NA
				APOM	NA	NA
	XXbac-BPG32J3			*C6orf47*	NA	NA
				BAT4	NA	NA
				CSNK2B	NA	NA
				LY6G5B	NA	NA
				LY6G5C	NA	NA
				LY6G6E	NA	NA
				BAT5	NA	NA
				LY6G6F	NA	NA
				LY6G6D	NA	NA
				LY6G6C	NA	NA
				C6orf25	NA	NA
				DDAH2	NA	NA
				CLIC1	NA	NA
				MSH5	0.4849	0.0027
				C6orf26	NA	NA
				C6orf27	NA	NA
				VARS	NA	NA
	XXbac-BPG254B15	6p21.33	G/L	*LSM2*	NA	NA
				HSPA1L	NA	NA
				HSPA1A	NA	NA
				HSPA1B	NA	NA
				C6orf48	NA	NA
				NEU1	NA	NA
				SLC44A4	NA	NA
				EHMT2	NA	NA
				C2	NA	NA
				ZBTB12	NA	NA
	XXbac-BPG116M5			*XXbac-BPG116M5.1*	NA	NA
				CFB	NA	NA
				RDBP	NA	NA
				SKIV2L	NA	NA
				DOM3Z	NA	NA
				STK19	NA	NA
				STK19P	NA	NA
				C4A	NA	NA
				AL645922.3	NA	NA
				TNXB	NA	NA
				CYP21A2	NA	NA
				TNXB	NA	NA
	XXbac-BPG15K13	6p21.33	G/L	*TNXA*	NA	NA

Table S1. Cont.

Region	Clone name	Band	Change	Candidate genes	Correlation coefficient	*P* value
	XXbac-BPG300A18	6p21.32	L	*CREBL1*	NA	NA
				AL662884.1	NA	NA
				FKBPL	NA	NA
				PRRT1	0.0665	0.6998
				PPT2	NA	NA
				EGFL8	NA	NA
				AGPAT1	NA	NA
				RNF5	NA	NA
				AGER	NA	NA
				PBX2	NA	NA
				GPSM3	NA	NA
				NOTCH4	NA	NA
	XXbac-BPG7F7			*C6orf10*	NA	NA
	XXbac-BPG495H8	6p21.32	L	*HLA-DRB9*	NA	NA
	XXbac-BPG161M6			*HLA-DRB5*	0.2444	0.1508
				HLA-DRB6	NA	NA
	XXbac-BPG254F23			*HLA-DRB1*	NA	NA
				HLA-DQA1	NA	NA
				AL662789.2	NA	NA
				HLA-DQB1	NA	NA
				HLA-DQA2	NA	NA
				HLA-DQB2	NA	NA
	XXbac-BPG226C15			*HLA-DQB3*	NA	NA
	XXbac-BPG186F10		G	*HLA-DMB*	NA	NA
				HLA-DMA	NA	NA
	XXbac-BPG181M17			*BRD2*	NA	NA
	XXbac-BPG185D15		L	*VPS52*	NA	NA
				RPS18	NA	NA
				B3GALT4	NA	NA
				WDR46	NA	NA
				PFDN6	NA	NA
				RGL2	NA	NA
				TAPBP	NA	NA
				ZBTB22	NA	NA
				DAXX	NA	NA
	XXbac-BPG294E21 RP4-570F3			*KIFC1*	NA	NA
				PHF1	NA	NA
				CUTA	NA	NA
				SYNGAP1	0.3443	0.0397
				ZBTB9	NA	NA
6	RP1-20C7	6p21.1	G/L	*PPP2R5D*	NA	NA
				MEA1	NA	NA
				KLHDC3	NA	NA
				C6orf153	NA	NA
				CUL7	NA	NA
				KLC4	NA	NA
	RP11-387M24		L	*CUL7*	NA	NA
				KLC4	0.3654	0.0284
				MRPL2	NA	NA
				PTK7	0.5467	0.0005
				SRF	NA	NA
	RP1-241K1		G	*No genes*	—	—
	RP1-302G2			*AL139392.1*	NA	NA
				SLC29A1	0.3230	0.0546
				HSP90AB1	NA	NA
				SLC35B2	NA	NA
				NFKBIE	NA	NA
				TMEM151B	NA	NA
				AL353588.1	NA	NA

Table S1. Cont.

Region	Clone name	Band	Change	Candidate genes	Correlation coefficient	*P* value
7	RP11-411F9	6q14.1	G/L	*No genes*	—	—
	RP11-328C7					
8	RP11-337M11	6q14.1	L	*AL132766.1*	NA	NA
	RP3-433F14					
9	RP1-101M23	6q21	L	*PRDM1*	0.1780	0.2989
	RP1-134E15			*ATG5*	−0.1	0.5616
10	RP1-281H8	6q25.1	L	*MAP3K7IP2*	0.0377	0.8270
				SUMO4	−0.2599	0.1257
				ZC3H12D	NA	NA
	RP11-703H16			*PPIL4*	0.0441	0.7985
	RP1-203A15			*LATS1*	−0.0447	0.7957
	RP11-52G20					
	RP1-317N9					
	RP11-350J20			*NUP43*	NA	NA
				PCMT1	0.1912	0.2639
				LRP11	0.2056	0.2289
				AL355312.3	NA	NA
				RAET1E	NA	NA
11	RP11-235G24	6q26	G/L	*No genes*	—	—
	RP3-428L16			*MAP3K4*	0.2502	0.1411
				MAP3K4	0.4099	0.013
	RP11-776J12			*AL139393.2*	NA	NA
				AL596452.1	NA	NA
	RP11-155H6					
	RP3-473J16			*AGPAT4*	0.1787	0.297
	KB-152G3			*PARK2*	0.3294	0.0499
	RP11-421L20					
	KB-521C9					
	RP1-45F6					
	RP1-119H20					
	KB-761D4					
	KB-1750A9					
	RP1-195L20					
	KB-611H5		L			
	KB-1954F7					
	KB-764F10					
	RP11-266I4					
	RP11-494F24					
	RP11-307K1					
	RP11-27G12					
	RP11-211O7			*PACRG*	0.2782	0.1004
	RP11-57O22		G/L			
	RP11-157L10					
	RP1-257A15			*AL078585.1*	NA	NA
	RP1-301L19					
	RP3-495O10					
	RP5-856H7			*QKI*	−0.0908	0.5985
	RP1-51J12					
12	RP11-471L1	6q27	G	*MLLT4*	0.3106	0.0652
	RP3-470B24					
	RP11-164L23			*C6orf123*	NA	NA
				KIF25	0.1397	0.4162
				FRMD1	0.3405	0.042
	RP11-3C9			*No genes*	—	—
	RP11-302L19			*No genes*	—	—

Table S2. Primers used for *PARK2* MSP PCR

Promoter status		Primer		No. nucleotides	Product size
Methylated	F5′	AGGTAAGTTTTTCGGTTGTTAAGCG	3′	25	178 bp
specific	R5′	GCGACCCAAAACCTACTAAAAATCG	3′	25	
Unmethylated	F5′	AAGTGATTGGTTAATATGGTGGGTG	3′	25	160 bp
specific	R5′	ACACAACCCAAAACCTACTAAAAATCAT	3′	28	