Supporting Information

Amos-Landgraf et al. 10.1073/pnas.1120753109

SI Materials and Methods

1. Model. Two probability models were used to analyze bivariate allelic ratios of mutant genomic DNA (gDNA) and mutant cDNA. The first model was a discrete Gaussian mixture model that provides for model-based clustering of the samples into three tumor classes. The class-specific mean parameters were constrained, and therefore, the expected gDNA value in the maintenance of heterozygosity of *adenomatous polyposis coli* (*Apc*) in gDNA and cDNA (gMOH/cMOH) class equaled the expected gDNA value in the gMOH/complementary loss of heterozygosity (cLOH) class; similarly, the expected cDNA values in classes gLOH/cLOH and gMOH/cLOH were constrained to be equal, which was justified by the measurement process. Otherwise, each class was allowed to have an unconstrained 2 × 2 covariance matrix. Owing to mean constraints, standard software could not be applied to fit the probability model; custom R code was developed to implement the expectation maximization algorithm and thus, estimate the means, covariances, and mixing proportions by the method of maximum likelihood. Data from normal tissues were included in the estimation but forced to arise from the gMOH/cMOH class.

A second model was developed to analyze the allelic ratios. This admixture model aimed to go farther than the discrete mixture model by representing each tumor as a mixture of cells of three different pure cell types. Computations in this admixture model were more complex than the computations required for the first model, and they are developed fully here. We analyzed the polyposis in the rat colon (Pirc) and Rb9 data separately. The n tissue samples from one experiment provided bivariate data, $\mathrm{DAT} = \{(X_i, Y_i): i = 1, 2, \ldots, n\}$, measuring mutant allele ratios in gDNA and cDNA, respectively. In the admixture model, the samples $\{i\}$ produced mutually independent data points. Furthermore, sample i was considered to be comprised of fractions of three pure cell types, with the fractions $U_{i,j} \geq 0$ and $\Sigma_{j=1}^{3} U_{i,j} = 1$. None of these fractions were observed, except for cases i that were normal tissue controls, for which we assumed $U_i = (U_{i,1}, U_{i,2}, U_{i,3}) = (1, 0, 0)$. For each tumor, the admixture vector U_i itself was viewed as a random draw from an unknown distribution $\pi(u)$ over the simplex $\mathcal{S}_3 = \{u = (u_1, u_2, u_3): u_j \geq 0, \sum_j u_j = 1\}$. Computations were enabled by taking a finite-grid approximation to $\mathcal{S}_3$ and thus, a vector approximation to π.

	Cell type	Fraction
1	gMOH/cMOH	$U_{i,1}$
2	gLOH/cLOH	$U_{i,2}$
3	gMOH/cLOH	$U_{i,3}$

We supposed that measurements of gDNA and cDNA on a pure tissue comprised only of cells of one type would be bivariate normal, with means and variances that depended on the cell type but without any correlation in the measurement error. The lack of correlation between gDNA and cDNA measurements on pure tissue was a key assumption that allowed us to infer admixture rates in the actual tissue samples. It was justified considering that pyrosequencing was performed separately for gDNA and cDNA material and that variation attributable to the initial isolation of all nucleic acids from the tissue was probably negligible. We also assumed that the distribution of data in one channel (gDNA or cDNA) was not affected by state of the other channel. In other words, for mean parameters a_1, a_2, b_1, and b_2, the mean gDNA measurement in pure gMOH cells was a_1, and the mean gDNA measurement in gLOH cells was a_2. Similarly the mean cDNA measurement in cMOH cells was b_1, and the mean cDNA measurement in cLOH cells was b_2. Fig. S4 shows these parameters as estimated from the Pirc dataset.

The contribution to gDNA/cDNA measurements (X_i, Y_i) from the pure cells of one type was assumed to be normally distributed, where the bivariate mean depended on the type and there was one variance parameter for gDNA and one variance parameter for cDNA, regardless of type. A key element was having independent measurement errors in these pure cells (the apparent marginal correlation in data arose from admixture). Thus, X_i and Y_i were treated as independent given admixture rates U_i, with gDNA data following (Eq. **S1**)

$$X_i|(U_i = u) \sim \mathrm{Normal}\{\mathrm{mean} = (u_1 + u_3)a_1 + u_2 a_2, \quad \mathrm{variance} = \sigma_a^2(u_1^2 + u_2^2 + u_3^2)\} \qquad \textbf{[S1]}$$

and cDNA data following (Eq. **S2**)

$$Y_i|(U_i = u) \sim \mathrm{Normal}\{\mathrm{mean} = u_1 b_1 + (u_2 + u_3)b_2, \quad \mathrm{variance} = \sigma_b^2(u_1^2 + u_2^2 + u_3^2)\}. \qquad \textbf{[S2]}$$

Conditional on the admixture rates, the unknown parameters were $\theta = (a_1, b_1, a_2, b_2, \sigma_a^2, \sigma_b^2)$.

2. Inference. The unknown objects were mean and variance parameters in the vector θ, admixture vectors $\{U_i = (U_{i,1}, U_{i,2}, U_{i,3})\}$, and the probability distribution $\pi(u)$ over the simplex $\mathcal{S}_3$. We placed a prior distribution on θ and $\pi(u)$ and developed Bayesian computations by Markov chain Monte Carlo to simulate the distribution of these unknowns conditional on data *DAT*. We approximated the simplex $\mathcal{S}_3$ by a finite grid of proportion vectors, denoted ugrid in the R code. This matrix was size 861 × 3 holding row vectors that represent possible realizations of each admixture vector U_i. K = 861 came by taking a regular 40 × 40 grid over the unit square, keeping coordinates for which the values sum to less than one and considering these two values to be the first two entries of a possible admixture vector (the third being one minus the sum). Thus, π was a length K = 861 vector holding the probability distribution governing the U_i values. In the Bayesian analysis, we placed a conjugate exchangeable Dirichlet prior over π, using a small prior mass α = 1 (**S3**):

$$\pi \sim \mathrm{Dirichlet}_K\left(\frac{\alpha}{K}, \frac{\alpha}{K}, \ldots, \frac{\alpha}{K}\right). \qquad \textbf{[S3]}$$

Next, we placed a noninformative flat prior on the mean parameters a_1, a_2, b_1, b_2 in θ, except that we insisted that all mean values exceed zero. We placed a weakly informative conjugate inverse-Gamma prior for the two variance parameters σ_a^2 and σ_b^2. We used a prior guess $\sigma_0^2 = 25$, with n_0 = 1. That is, the prior for each inverse variance was Gamma with shape $n_0/2$ and rate $n_0\sigma_0^2/2$. Evidently, the posterior information was relatively

high, and we would have computed very similar estimates for a wide range of prior hyperparameters.

We developed a systematic scan Gibbs sampler with separate updates for the means and variances in θ, the admixture vectors, and the mixing proportions π. (The following three subsections describe details of the Gibbs sampler.) After preliminary testing, we ran the sampler for 500,000 scans, saving 1 in each 100 states and basing posterior estimates on 5,000 putative draws from the joint posterior distribution. Output analysis indicated very good mixing of the Markov chain. Mean values of the output were used for parameter estimation.

2.1. Updating θ. The gDNA and cDNA measurements X_i and Y_i were conditionally independent given the admixture vector U_i. The parameters a_1, a_2, σ_a^2 refer to the distribution of X_i, and b_1, b_2, σ_b^2 refer to the distribution of Y_i. Considering the similarity of gDNA and cDNA models, we show here only the Gibbs sample update rules for the parameters governing X_i. Re-expressing the observation model from the previous section, we have (**S4**)

$$X_i|\text{else} \sim \text{Normal}\left(c_i^T\mu, d_i\sigma_a^2\right), \qquad \textbf{[S4]}$$

where $d_i = U_{i,1}^2 + U_{i,2}^2 + U_{i,3}^2$ multiplied by variance $\mu = (a_1, a_2)^T$ holds the pure cell means, and (Eq. **S5**)

$$c_i = \begin{pmatrix} c_{i,1} \\ c_{i,1} \end{pmatrix} = \begin{pmatrix} U_{i,1} + U_{i,3} \\ U_{i,2} \end{pmatrix}. \qquad \textbf{[S5]}$$

(In both c_i and d_i, the roles of $U_{i,1}$ and $U_{i,2}$ are reversed when considering the cDNA data Y_i.) Using standard Bayesian arguments and a flat prior, we find the full-conditional distribution for the mean parameters a_1 and a_2 is (**S6**)

$$\begin{pmatrix} a_1 \\ a_1 \end{pmatrix}\Bigg|\text{else} \sim \text{Normal}\left\{\begin{pmatrix} m_1 \\ m_2 \end{pmatrix}, \begin{bmatrix} s_{1,1} & s_{1,2} \\ s_{1,2} & s_{2,2} \end{bmatrix}\right\}, \qquad \textbf{[S6]}$$

where posterior means are (Eq. **S7**)

$$m_1 = \frac{1}{1 - \frac{C^2}{B_1B_2}}\left(\frac{A_1}{B_1} - \frac{A_2C}{B_1B_2}\right) \qquad \textbf{[S7]}$$

and (Eq. **S8**)

$$m_2 = \frac{1}{1 - \frac{C^2}{B_1B_2}}\left(\frac{A_2}{B_2} - \frac{A_1C}{B_1B_2}\right). \qquad \textbf{[S8]}$$

The posterior covariance matrix is (Eq. **S9**)

$$\begin{bmatrix} s_{1,1} & s_{1,2} \\ s_{1,2} & s_{2,2} \end{bmatrix} = \frac{1}{B_1B_2 - C^2}\begin{bmatrix} B_2 & -C \\ -C & B_1 \end{bmatrix}. \qquad \textbf{[S9]}$$

The contributing quantities are (Eq. **S10**)

$$A_1 = \sum_{i=1}^{n} \frac{1}{d_i\sigma_a^2} x_i c_{i,1}, \qquad \textbf{[S10]}$$

(Eq. **S11**)

$$A_2 = \sum_{i=1}^{n} \frac{1}{d_i\sigma_a^2} x_i c_{i,2}, \qquad \textbf{[S11]}$$

(Eq. **S12**)

$$B_1 = \sum_{i=1}^{n} \frac{1}{d_i\sigma_a^2} c_{i,1}^2, \qquad \textbf{[S12]}$$

(Eq. **S13**)

$$B_2 = \sum_{i=1}^{n} \frac{1}{d_i\sigma_a^2} c_{i,2}^2, \qquad \textbf{[S13]}$$

and (Eq. **S14**)

$$C = \sum_{i=1}^{n} \frac{1}{d_i\sigma_a^2} c_{i,1}c_{i,2}. \qquad \textbf{[S14]}$$

Thus, the Gibbs update of a_1 and a_2 arose from the bivariate normal posterior given in expression **S6**. We sampled it by sampling the marginal of a_1 and then, the induced conditional of a_2 given a_1. We imposed the constraint of $a_1, a_2 > 0$ in both updates.

Under the conjugate inverse-Gamma prior indicated above, the variance parameter σ_a^2 has an inverse-Gamma full-conditional distribution. More specifically (Eq. **S15**),

$$\frac{1}{\sigma_a^2}\Bigg|\text{else} \sim \text{Gamma}\left\{\text{shape} = \frac{n_0 + n}{2}, \text{rate} = \frac{1}{2}\left(n_0\sigma_0^2 + \sum_{i=1}^{n}\frac{\left(x_i - c_i^T\mu\right)^2}{d_i}\right)\right\}. \qquad \textbf{[S15]}$$

Analogous updates for the Y_i parameters b_1, b_2, and σ_b^2 are as above but with the roles of $U_{i,1}$ and $U_{i,2}$ reversed.

2.2. Updating admixture vectors. Tumor i has admixture vector U_i, which has a complicated but discrete conditional distribution over the K possible rows in ugrid given the data and the parameters θ and $\pi = (\pi_u)$. For a row u of ugrid, we have (Eq. **S16**)

$$P(U_i = u|\text{DAT}, \theta, \pi) \propto p(x_i|u, \theta)p(y_i|u, \theta)\pi_u. \qquad \textbf{[S16]}$$

In our implementation, we find the logarithm of the right-hand side by invoking the normal model from *SI Materials and Methods, 2.1.* and the estimated sampling model for U_i. We renormalize for each i to get the full conditional distribution for each U_i and then run the Gibbs update by sampling these discrete distributions one time each in parallel.

2.3. Updating mixing proportions. Given everything else, the vector π depends only on how many admixture vectors take each of the possible values in ugrid. By conjugacy of the Dirichlet relative to these multinomial counts (**S17**),

$$\pi|\text{else} \sim \text{Dirichlet}_K\left(\frac{\alpha}{K} + s_1, \frac{\alpha}{K} + s_2, \ldots, \frac{\alpha}{K} + s_K\right), \qquad \textbf{[S17]}$$

where s_j counts how many U_i values take value ugrid[j,]. This sample is by renormalizing independent and properly Gamma-distributed variables.

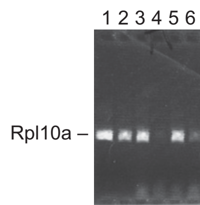

Fig. S1. RT-PCR for the ribonuclear protein subunit *Rpl10a* in mouse Rb9-*Multiple intestinal neoplasia* (*Min*) tumors to assess cDNA qualitatively. Lanes 1 and 2 are tumors that maintained heterozygosity of *Apc* in the DNA fraction but were below the level of detection by pyrosequencing of cDNA for *Apc*. Lanes 3–6 are other mouse tumors, both positive and negative for cDNA.

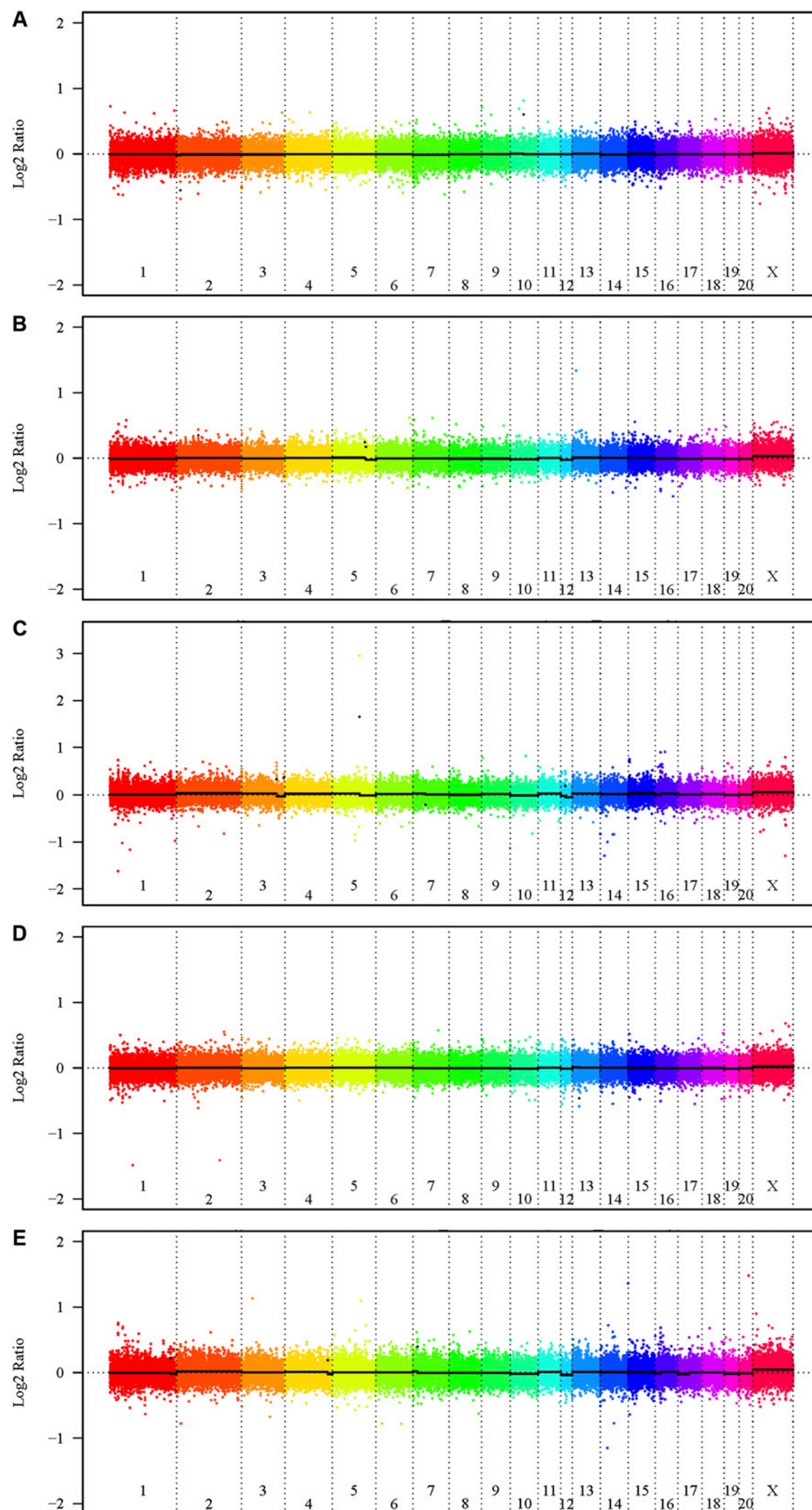

Fig. S2. Array comparative genomic hybridization (aCGH) was performed using Nimblegen rat whole-genome arrays and two-color labeling (Cy5 and Cy3) with tumor and normal spleen DNA. Two male F344-Pirc animals and one age-matched male F344/Tac WT rat were killed at 9 mo of age. From each F344-Pirc animal (P1 and P2), DNA was isolated immediately after necropsy, starting with 50 mg from the largest single colonic tumor and spleen. DNA was extracted using the Qiagen DNeasy tissue kit. Comparative genomic hybridizations were performed by Nimblegen Systems in their manufacturing facilities in Iceland using the Nimblegen RGSC 3.4 isothermal rat aCGH chip with 385,000 unique sequence features and a median probe density of 5,303 bp. One aliquot of tumor (T) and spleen (S) DNA from each animal was labeled with Cy5, and a second aliquot was labeled with Cy3. Spleen DNA sample P1S was hybridized against WT spleen DNA (*A*). Reciprocal hybridizations of Cy5 and Cy3 were performed: P1T vs. P1S (*B* and *C*). Similarly, Reciprocal hybridizations of P2T and P2S were performed (*D* and *E*). aCGH plots were generated using NimbleScan software, and the data were analyzed using SignalMap v1.8.

Fig. S3. LOH analysis of F1-Pirc tumors for markers on chromosomes 4, 7, 10, 15, 16, and 18q, not including markers for chromosome 1 (Table S2) using SNPs listed in Table S1. Each point represents a tumor. Not all tumors were tested for all markers. No significant loss of the F344 or ACI alleles was seen in any tumors.

Fig. S4. Bivariate pyrosequencing data from normal tissue (open circles) and Pirc tumors (filled circles). Components of the fitted admixture model are shown. Pure cells of each of three types would yield bivariate Gaussian measurements from the three bivariate normal distributions (contours show regions holding 95% probability in the fitted component). The means are constrained (gray lines). Each tumor sample is viewed as an unknown mixture of the three cell types.

Table S1. LOH analysis of gDNA from F1-Pirc tumors using quantitative allele-specific pyrosequencing

			Contribution of F344 allele (%)	
Marker	Chr	Position (Mb)	Tumor X	Other tumors (mean ± SD)
SNP-76	1p	7	48.1	49.3 ± 3.0
SNP-102	1q	46	47.8	ND
SNP-82	1q	131	**84.7**	53.4 ± 0.6
SNP-75	1q	159	**78.7**	49.5 ± 1.4
SNP-83	1q	232	**78.3**	43.5 ± 1.8
SNP-2	1q	233	**85.8**	54.8 ± 3.1
Pirc	18p	27	**85.0**	83.5 ± 15.1*
SNP-33	18q	83	49.4	51.2 ± 3.3

SNPs across the genome were used to detect significant deviations from the expected 50% contribution from each of the parental strains. Tumor X showed LOH on Chr 1q, with a cross-over between 46 and 131 Mb. Allele ratios in bold show gLOH on one arm of the salient chromosome. Seven other tumors from the same animal serve as controls. Chr, chromosome; ND, not done.

*Six tumors were evaluated for Pirc.

Table S2. Mutations to *Apc* in gMOH/cMOH F1-Pirc tumors

Mutation type	Position on Chr. 18	Position in *Apc*	Treatment
T to C	26,782,134	Intron 13	DSS
C to T	26,783,847	Exon 15	None
C to T	26,784,615	Exon 15	DSS

The mutation type and base pair position (rat genome build Baylor 3.4, November 2004) of the three identified *Apc* mutations of the 12 gMOH tumors sequenced. Chr, chromosome; DSS, dextran sulfate sodium.

Table S3. Admixture probabilities for Rb9-Min and F1-Pirc tumors

Model	Tumor category	Probability (majority)	Probability (plurality)
Rb9-Min	gMOH/cMOH	0.33	0.45
Rb9-Min	gLOH/cLOH	0.43	0.51
Rb9-Min	gMOH/cLOH	0.02	0.04
F1-Pirc	gMOH/cMOH	0.19	0.19
F1-Pirc	gLOH/cLOH	0.63	0.68
F1-Pirc	gMOH/cLOH	0.12	0.13

The admixture probabilities for the majority or plurality of cells of each of the three tumor cell classes present in a given sample for both Rb9-Min and F1-Pirc as described in *SI Materials and Methods*.

Table S4. Primers

Primer name	Location	Forward	Reverse	Sequencing	SNP with surrounding sequence
Apc Alt-F	chr18: 26,777,147 + 26,777,166	GCAAGTGTTTTGCGGAATTT	NA	NA	NA
Apc Alt-R1	chr18: 26,782,136 + 26,782,155	NA	AGAGTTGATTCCTATTCATG	NA	NA
Apc Alt-R2	chr18: 26,782,307 + 26,782,326	NA	ACATTCCGTAATATCCCACC	NA	NA
APC Alt Promoter	chr18: 26,720,834 + 26,721,818	CCAGTCACTCGGTTCTGTCA	TACCACTGCCTGTGTCCTTG	NA	NA
Apc Promoter	chr18: 26,731,661 + 267,32,322	GAATGGCTGGAGTTCTTTGG	TGCCTCAGTTTCCAGTTCTGT	NA	NA
Apc Exon 1	chr18: 26,732,010 + 26,732,410	CCCATGTTCCGTCCTCTTAG	TGTCTTAGCTATCCCAAGTTTACAA	NA	NA
Apc Exon 2	chr18: 26,739,223 + 26,739,623	TGGCATGTTGTCAGTGTGAA	GATCTGAGGGGAAACGTCAA	NA	NA
Apc Exon 3	chr18: 26,739,951 + 26,740,394	CTGACTTTGCAGGGCAAGTT	CAGGCTGCCACCTATCATCT	NA	NA
Apc Exon4	chr18: 26,742,943 + 26,743,399	TAGCAGGGAAGGAGCACACT	AATGACGACGGTGACACTCA	NA	NA
Apc Exon 5	chr18: 26,747,107 + 26,747,506	TCATAGTTTAGGATTCCTTGGAAGA	TTTGATGGTGCTGCAGGTTA	NA	NA
Apc Exon 6	chr18: 26,751,039 + 26,751,389	TTGGAAAGCAGAAGCCAAAG	CGCATGTAACAAAAGGGGACT	NA	NA
Apc Exon 7	chr18: 26,762,128 + 26,762,608	TGAGAGGAGCACATCTGGTG	TCCGGAGCATAAAACATTCC	NA	NA
Apc Exon 8	chr18: 26,766,510 + 26,766,983	CTGCAGTCTGCTGGTTGTGT	GTCTGGAGCTTAGGCTGCTC	NA	NA
Apc Exon 9	chr18: 26,768,835 + 26,769,655	TGACCTTTAACTGATGCTGACAA	TTTTTCAGGGAAATGCAGATT	NA	NA
Apc Exon 10	chr18: 26,770,811 + 26,771,210	ACTGCTGGCTGTTCTCCTGT	ATATGGATGGAAGCAGCTGAA	NA	NA
Apc Exon 11	chr18: 26,775,431 + 26,775,923	CTGGCAGGGAGTGAAATTGT	CAGTGTGAGAAAGGGGCTTC	NA	NA
Apc Exon 12	chr18: 26,776,233 + 26,776,784	TTTCATTTGGGCAGTTTGTTC	TTGGGTTAGGAATGCTCCAC	NA	NA
Apc Exon 13	chr18: 26,776,833 + 26,777,287	GGGAAAACCTATGTTCTGAAGC	CTATTCTCGTACCCCACATTTT	NA	NA
Apc Exon 14	chr18: 26,782,006 + 26,782,505	TGCATCATGGAAAGATACGTG	TGGAGTTGATAGTGGTCTTCACA	NA	NA
Apc Exon 15-1	chr18: 26,783,492 + 26,784,011	ACGATTGCGATGTGAACGTA	CCCCATAGCGATCATTTTGT	NA	NA
Apc Exon 15-2	chr18: 26,783,179 + 26,784,469	TGCTCACAAGTGCTTTGGAC	GAAGAGGTTCCCGGATTTTC	NA	NA
Apc Exon 15-3	chr18: 26,784,077 + 26,785,123	GTCTCCTGGCTCAAGTTTGC	TGCTGTTGGAACTTGAGGTG	NA	NA
Apc Exon 15-4	chr18: 26,785,044 + 26,786,394	AGGCAATCAAGAAGCCAGAA	AGGGCTCATCCAGACTCAGA	NA	NA
Apc Exon 15-5	chr18: 26,786,008 + 26,787,066	CTTCTGTCAGCTCCCTGGAC	GCTTGTGGCTTCTTCCTTTG	NA	NA
Apc Exon 15-6	chr18: 26,787,047 + 26,788,273	CAAAGGAAGAAGCCACAAGC	GACTTCAGGGACAGGATGGA	NA	NA
Apc Exon 15-7	chr18: 26,788,000 + 26,789,006	AGAGGAGGCCTTCCAGACTC	CAAGCCTGTTTGTTTGCTGA	NA	NA
Apc Exon 15-8	chr18: 26,788,977 + 26,790,026	CAGCAAAACCTCAGCAAACA	AAAGGAGTTCAGGCGGTTTT	NA	NA
Apc Exon 15-9	chr18: 26,789,791 + 26,790,663	AGATGGCACCTGCTGTTTCT	AAAGGACAATGGAACGGATG	NA	NA
SNP-76	chr1: 7,422,419 – 7,422,495	Biotin-TCCTCCAAAGTGTTCAACCA	GAGGAGGTCTGGTCGTGTTG	CTTCACTATCTGGATCTG	GAAACC[G/T]CGTGGTTGAA CACTT
SNP-102	chr1: 45,963,406 – 45,963,533	Biotin-ATGAGGGCTCAAAAA AATTACCTA	GTGTGCTTATTATTTGGGGCTATG	CTGTTATTTCCTTCTACCAT	G[C/T]AGATTCAGGGGATGAATT CATGTCA
SNP-82	chr1: 131,354,285 – 131,354,366	Biotin-GGGAGCAGGAAC AACTGAAG	CTCTTGGGTACCTGGCAAAG	TCCATCTGCCTCTGGCTG	AGC[T/C]GCTCTGCATCCCTCT
SNP-75	chr1: 158,844,845 + 158,844,927	Biotin-AAGAGAGGCGG TTCTCAGGT	ACAGAGAGCCGGGAGGAG	GAGGGCAGCGATGACGAC	ACCTC[G/A]CTCACCTGAGAA CCGCCT
SNP-83	chr1: 231,882,065 – 231,882,156	Biotin-CCACTGCGAAG AGATTCAC	AGTGCGGAACTCACTGTCTG	TGCCTTCAGCTTCCTCGA	GCCTCTT[T/C]GGTGAATCTCTTCG
SNP-2	chr1: 232,855,852 – 232,855,932	Biotin-ACCCCCATTGTTGT GTGACT	CCTGAGACGGAAAAACTGGA	AGGTAACTCTGGAAGGGA	GTGACTGCCCCCCAT[T/C]CTCC CTTCC
SNP-58	chr4: 75,039,070 + 75,039,137	CCAAGTCTGCTACTTCCCTCA	Biotin-GAACATCCATGCGACATCC	AGGGGCGACTCAGTGCTT	ACTCAGTGCTTT[C/G]TTTCCA AGCGGAT

Table S4. Cont.

Primer name	Location	Forward	Reverse	Sequencing	SNP with surrounding sequence
SNP-61	chr7: 71,279,026 + 71,279,107	GCATGCCCTGTATCAAGACC	Biotin-GCTTCCTTTCTTA TGTCTTGTTGG	GACCATGTATGCAGTGCT	ATGTATGCAGTGCTA[C/G]TGTC CTGGAGT
SNP-13	chr7: 128,110,496 − 128,110,607	Biotin-TGGCTAAGAAGTGAA TCACGAA	GCAGAGCCTCAGGCAATTTA	TTCATCTTCTTTGGGGTA	CAATGACTTAGA[A/G]GTACCCC AAAGAA
SNP-65	chr10: 25,359,997 + 25,360,102	Biotin-CCTTCAAGAGCTC CAAGCTG	GGGAATGAGAAGGGAGTTGA	TTGTGAACTAAAAGACAT	CTTCACTACCCC[A/G]TATGTCTTTT
SNP-27	chr15: 13,867,393 − 13,867,502	Biotin-ATTTGCTGGAATGG CTTCTC	GGAGGTTTCCTAATGGAGCTG	GAATTAAAGGACAGGTAA	CTTTCTTCTCTGAATCAC[A/G]C TTACCTG
SNP-28	chr16: 9,709,784 + 9,709,903	Biotin-GCCAAATACAACGC AGACCT	GGCATTGCCAGACTCTTGAC	GGACAGACTTGGCTGAAA	GTGCTCCATACTCAG[A/G]GATGT ACTTG
SNP-33	chr18: 838,18,221 − 83,818,328	Biotin-AGGCAAGGGA TGATTCTCCT	TGTCACTTGCCAATGAGGAC	AAAAGTTCCAGCTGTCAG	CCAGCCTGTCAGA[T/A]GGCCCCT AGGAG
Apc Min (Mouse)	chr18: 34,472,100 + 34,472,293	TTTTGACGCCAATCGACATG	Biotin-GATGGTAAGCACTGAG GCCAATA	CGTTCTGAGAAAGACA GAAG	
Apc Pirc (Rat)	chr18: 26,785,219 + 26,785,307	ATGTGAACCAGTCTTTGTG TCAG	Biotin- ATGCTGTTCTTCCTCAG AATAACG	GGAAGACGACTATGAAGAT	

The list of PCR primers used for either the mutation detection in tumors or confirmation and allele-specific pyrosequencing of SNPs. Alt primers were used to confirm the alternative splicing between exons 13 and 14 (Fig. 2). Chr, chromosome.

Other Supporting Information Files

Dataset S1 (PDF)