Elevated microsatellite alterations with allelic loss on 9p24.2 signify less aggressive metastasis in colorectal cancer

Minoru Koi,[1, 2] Melissa Garcia,[2] Chan Choi,[3] Hyeong-Rok Kim,[4] Junichi Koike,[5] Hiromichi Hemmi,[6] Takeshi Nagasaka,[7] Yoshinaga Okugawa, [2, 8] Yuji Toiyama,[8] Takahito Kitajima,[8] Hiroki Imaoka[8], Masato Kusunoki[8], Yin-Hsiu Chen,[9] Bhramar Mukherjee,[9] C Richard Boland, [2] John M Carethers[1]

Supplementary Materials and Methods

KRAS/BRAF Mutation Detection

The following primers were used for pyrosequencing : *KRAS* forward primer:

5'-GGCCTGCTGAAA ATGACTGA-3', *KRAS* reverse primer:

5'-BioATAGCCTCAATTCTTACCATCC-3', *KRAS* sequencing primer:

5'-AGGTGATTTTGGTCTAGCTACAG-3', *BRAF* forward primer:

5'-GGCCTGCTGAAAATGACTGA-3', *BRAF*-reverse primer: 5'-Bio-TAGCTGTATCGT

CAAGGCACTCT-3' and *BRAF* sequencing primer: 5'-TTGTGGTAGTTGGAGCT-3'.

Data mining and screening for a gene associated with E/L in LM

In total, we selected 141 genes with di-, tri- or tetranucleotide repeats for screening. The main criteria for selection of these genes were: 1) microsatellite repeats were at 5'-UTR, exon, 3'UTR or intron of a gene, 2) the repeats were large enough to be susceptible for DNA polymerase error, and 3) the repeats were polymorphic in length so that LOH can be detected. An NCBI blast search (http://blast.ncbi.nlm.nih.gov/Blast.cgi) followed by accessing the Ensemble Database (http://www.ensembl.org/index.html) to detect polymorphism identified 24 genes with polymorphic tetranucleotide repeats. For selecting genes with trinucleotide repeats, we used a database published by Kozlowski et al, where 878 genes with more than 6 units of trinucleotide repeats are listed (3). Among them, we selected 64 polymorphic genes with more than 8 units of trinucleotide repeats in their 5'-UTR, an exon, or 3'-UTR. To select a

gene with dinucleotide repeats, we used the Satellog Database (4). We selected 44 genes containing polymorphic CA/GT with more than 8 units (8-49 units) in their 5'UTR, an intron or 3'UTR (Table S2). The association of a selected gene with "cancer" was examined by accessing the NCBI PubMed literature database (http://www.ncbi.nlm.nih.gov/pubmed). Seventy percent of the selected genes have been reported to be associated with cancer in literature. In addition to the above genes, we added the 9 EMAST markers that were frequently mutated in cancer tissues to the list (5). Template DNA from 24 cases of LM tissues that exhibited MSI-M and matched normal tissues were amplified for each of these loci and analyzed for FS and LOH.

DNA Isolation, FS and LOH Analysis

For DNA extraction, tumor and normal tissues were micro-dissected separately from formalin-fixed paraffin-embedded (FFPE) 10μm sections. Genomic DNA was isolated and purified from micro-dissected tissues using a QIAamp DNA FFPE Tissue purification kit (QIAGEN, Valencia, CA).

For 141 gene loci containing polymorphic di-, tri- or tetranucleotide repeats, the genomic sequences from both 5' and 3' ends of the repeats were used to design PCR primers by the online software Prim3Plus (http://www.bioinformatics.nl/cgibin/primer3plus/primer3plus.cgi). Amplification of these loci and detection of FS or LOH were performed using the method of Schuelke (2). After heat denaturation, amplified PCR products were electrophoresed on an ABI PRISM 3100 Avant Genetic Analyzer (Applied Biosystems, Foster City, CA) and analyzed by GeneMapper fragment analysis software (Applied Biosystems).

A locus was determined FS-positive when a PCR product generated from tumor tissue exhibited at least one new peak compared to the product from the matching normal tissue. When a normal tissue exhibited heterozygosity at a particular marker, LOH was assessed in the corresponding tumor tissue. The height of the electropherographic peak of the PCR product was used as a measure for signal intensity. The ratio of signal intensities between the two alleles in normal tissue, and the ratio of signal intensities between two alleles in the corresponding tumor were compared. When the ratio in tumor was less than 50% of the ratio in normal, the locus was determined to be LOH positive.

***9p24.2* LOH**

Seven polymorphic markers: *2-4*, *505*, *6160*, *6640*, *6060*, *230*, and *240* were used to detect LOH from an approximately 370 Kb region spanning the *SMARCA2* locus. We defined 9p24.2 LOH as positive when one of the seven markers showed LOH, and as not informative when homozygosity was detected in all seven markers. The remainder was defined as 9p24.2 LOH negative. The primer sequences for these loci are as follows: *2-4*-F (5'-AGCCTGAACACTGCATAGTGA-3', *2-4*-R (5'- TCATCTTTTGGAAATGGAATAAGG -3'), *505*-F (5'-AGCTGTGCTAGCGCCTACTC-3'), *505*-R (5'-GGCTGAGAGCAGAAGAATGG-3'), *6160*-F (5'-GTGAAACCCATTCTTGCTAC-3'), *6160*-R (5'-ATCCTGACACTGAACTCCAG-3'), *6640*-F (5'-ACGATAAAGCAAGTACAGCAA-3'), *6640*-R (5'-AGCCTGGACAACAAGAGTGAA-3'), *6060*-F (5'-TTGTAACCACCTGAAGTAGAAGAG-3'), *6060*-R (5'-CAGCTACTCCCAGGCTTGAA-3'), *230*-F (5'-GAAACATAACCAAGAAGATGGATG-3'), *230*-R (5'-CTTCTGCAATGGTGTAGCCAGT-3'), *240*-F 5'-TTTTTAAACAGCCCAACTTTCA-3') and *240*-R (5'- TCACACCCACTTTTCAGAGG-3'). We defined LOH as positive when one of the four markers showed LOH, and as not informative when

homozygousity was detected in all three markers. The remainder of the cases was defined as non-LOH.

Statistical Analysis

Because Tumor set #4 is more biased to recurrent cases, a probability weight for each case was determined before analysis. Probability weight was calculated based on the recurrence rate after surgery. Recurrence rate at DPCNU was 12.1%, 10%, 3.3%, 0.6% and 1.3% at the first, second, third, fourth and fifth year after surgery respectively. Therefore, the expected number of recurrence for each preceding year among 124 cases was 15, 12.4, 4.092, 0.744, 1.612 and expected number of non-recurrence cases was 90.148 while observed number of recurrence cases was 25, 17, 6, 8, and 1 cases and non-recurrence number was 67. Because probability weight was a ratio between number of expected case and number of observed cases, 0.6, 0.729, 0.682, 0.093, and 1.612 were obtained for the recurrence at the first, second, third, fourth and fifth year after surgery respectively. Probability weight for non-recurrence was 1.345.