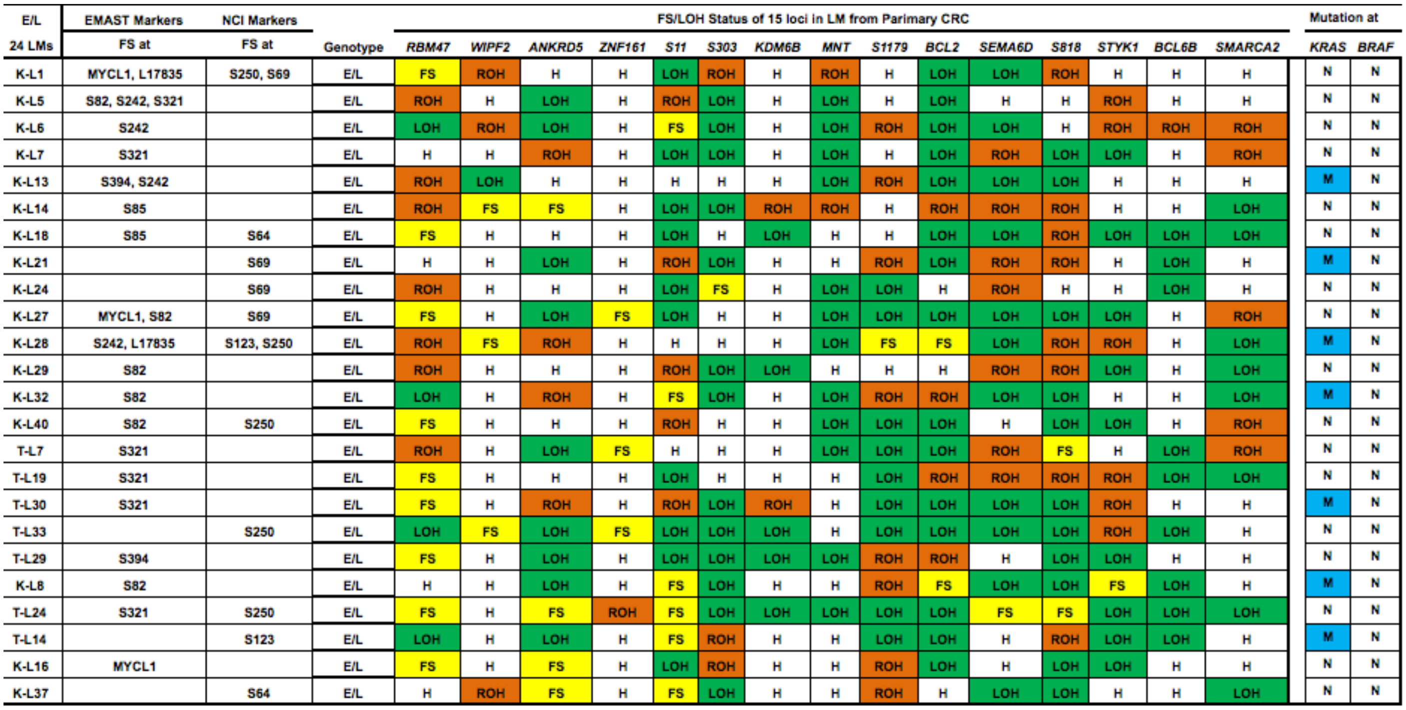
SUPPLEMENTARY MATERIALS: FIGURES

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 FIGURES AND FIGURE LEGENDS

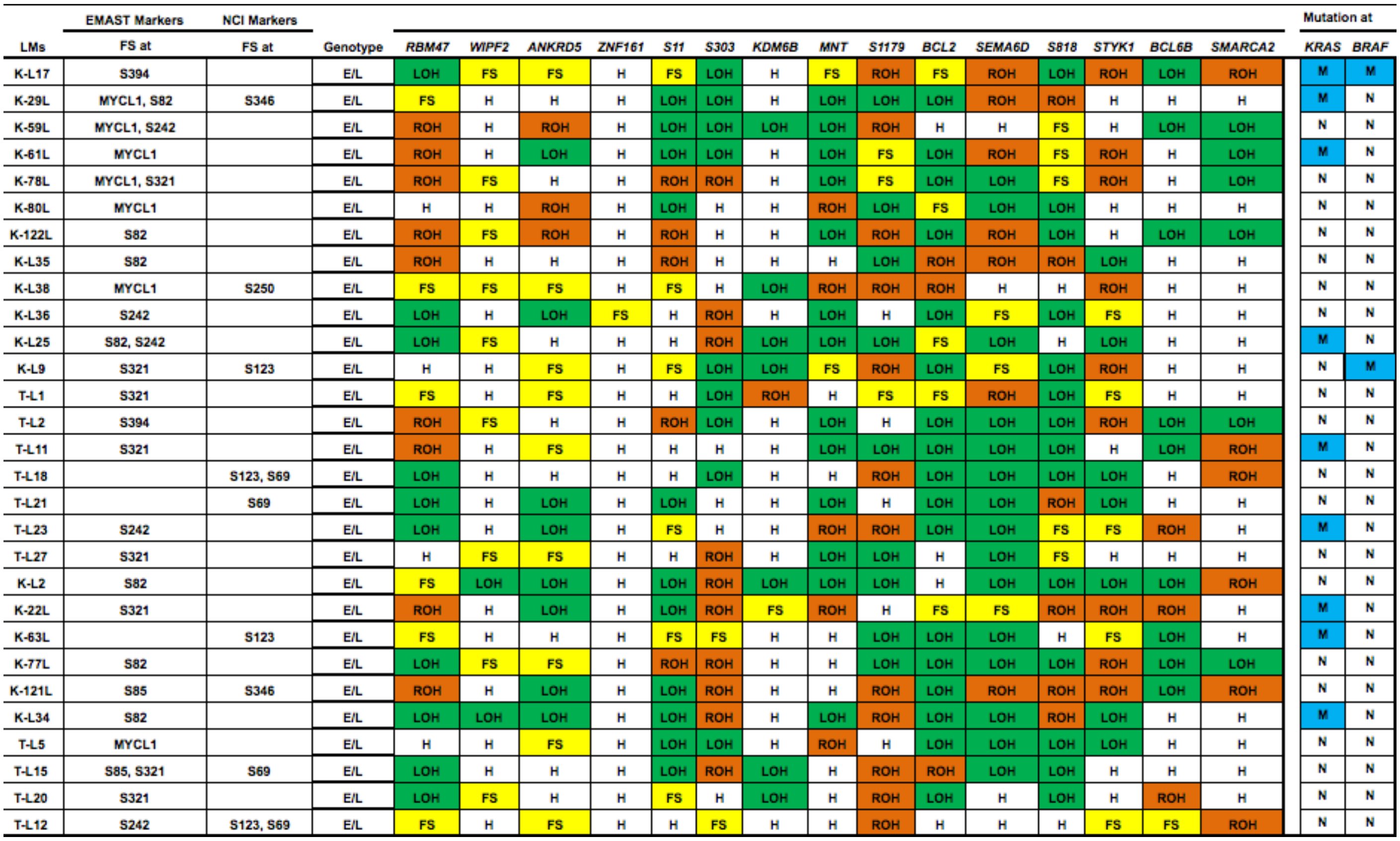
FIGURE 1A.



E/L	EMAST Markers	NCI Markers		FS/LOH Status of 15 loci in LM from Parimary CRC															Mutation at	
24 LMs	FS at	FS at	Genotype	*RBM47*	*WIPF2*	*ANKRD5*	*ZNF161*	*S11*	*S303*	*KDM6B*	*MNT*	*S1179*	*BCL2*	*SEMA6D*	*S818*	*STYK1*	*BCL6B*	*SMARCA2*	*KRAS*	*BRAF*
K-L1	MYCL1, L17835	S250, S69	E/L	FS	ROH	H	H	LOH	ROH	H	ROH	H	LOH	LOH	ROH	H	H	H	N	N
K-L5	S82, S242, S321		E/L	ROH	H	LOH	H	ROH	LOH	H	LOH	H	LOH	H	H	ROH	H	H	N	N
K-L6	S242		E/L	LOH	ROH	LOH	H	FS	LOH	H	LOH	ROH	LOH	LOH	H	ROH	ROH	ROH	N	N
K-L7	S321		E/L	H	H	ROH	H	LOH	LOH	H	LOH	H	LOH	ROH	LOH	LOH	H	ROH	N	N
K-L13	S394, S242		E/L	ROH	LOH	H	H	H	H	H	LOH	ROH	LOH	LOH	LOH	H	H	H	M	N
K-L14	S85		E/L	ROH	FS	FS	H	LOH	LOH	ROH	ROH	H	ROH	ROH	ROH	H	H	LOH	N	N
K-L18	S85	S64	E/L	FS	H	H	H	LOH	H	LOH	H	H	LOH	LOH	ROH	LOH	LOH	LOH	N	N
K-L21		S69	E/L	H	H	LOH	H	ROH	LOH	H	H	ROH	LOH	ROH	ROH	H	LOH	H	M	N
K-L24		S69	E/L	ROH	H	H	H	LOH	FS	H	LOH	LOH	H	ROH	H	H	LOH	H	N	N
K-L27	MYCL1, S82	S69	E/L	FS	H	LOH	FS	LOH	H	H	LOH	LOH	LOH	LOH	LOH	LOH	H	ROH	N	N
K-L28	S242, L17835	S123, S250	E/L	ROH	FS	ROH	H	H	H	H	LOH	FS	FS	LOH	ROH	ROH	H	LOH	M	N
K-L29	S82		E/L	ROH	H	H	H	ROH	LOH	LOH	H	H	H	ROH	ROH	LOH	H	LOH	N	N
K-L32	S82		E/L	LOH	H	ROH	H	FS	LOH	H	LOH	ROH	ROH	LOH	LOH	H	H	LOH	M	N
K-L40	S82	S250	E/L	FS	H	H	H	ROH	H	H	LOH	LOH	LOH	H	LOH	LOH	H	ROH	N	N
T-L7	S321		E/L	ROH	H	LOH	FS	H	H	H	LOH	LOH	LOH	ROH	FS	H	LOH	ROH	N	N
T-L19	S321		E/L	FS	H	H	H	LOH	H	H	H	LOH	ROH	ROH	ROH	ROH	LOH	LOH	N	N
T-L30	S321		E/L	FS	H	ROH	H	ROH	LOH	ROH	H	LOH	LOH	LOH	LOH	ROH	H	H	M	N
T-L33		S250	E/L	LOH	FS	LOH	FS	LOH	LOH	LOH	H	LOH	LOH	LOH	LOH	ROH	LOH	H	N	N
T-L29	S394		E/L	FS	H	LOH	H	LOH	LOH	LOH	LOH	ROH	ROH	H	LOH	LOH	H	H	N	N
K-L8	S82		E/L	H	H	LOH	H	FS	LOH	H	H	ROH	FS	LOH	LOH	FS	LOH	H	M	N
T-L24	S321	S250	E/L	FS	H	FS	ROH	FS	LOH	LOH	LOH	LOH	LOH	FS	FS	LOH	LOH	LOH	N	N
T-L14		S123	E/L	LOH	H	LOH	H	FS	ROH	H	H	LOH	LOH	H	ROH	LOH	LOH	H	M	N
K-L16	MYCL1		E/L	FS	H	FS	H	LOH	ROH	H	H	ROH	LOH	H	LOH	LOH	H	H	N	N
K-L37		S64	E/L	H	ROH	FS	H	FS	LOH	H	H	ROH	H	LOH	LOH	H	H	LOH	N	N

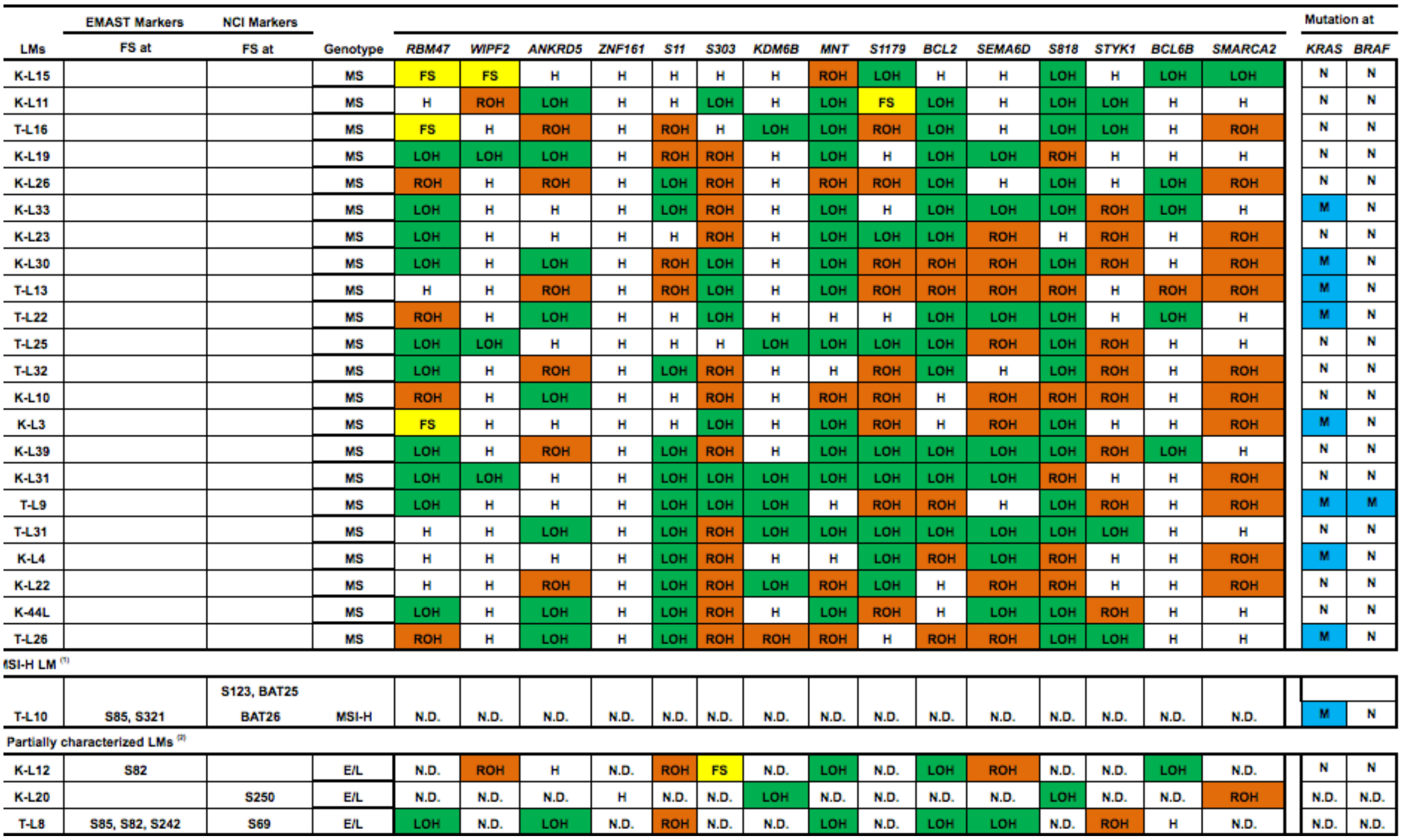
SUPPLEMENTARY MATERIALS: FIGURES

FIGURE 1B.



	EMAST Markers	NCI Markers																	Mutation at	
LMs	FS at	FS at	Genotype	*RBM47*	*WIPF2*	*ANKRD5*	*ZNF161*	*S11*	*S303*	*KDM6B*	*MNT*	*S1179*	*BCL2*	*SEMA6D*	*S818*	*STYK1*	*BCL6B*	*SMARCA2*	*KRAS*	*BRAF*
K-L17	S394		E/L	LOH	FS	FS	H	FS	LOH	H	FS	ROH	FS	ROH	LOH	ROH	LOH	ROH	M	M
K-29L	MYCL1, S82	S346	E/L	FS	H	H	H	LOH	LOH	H	LOH	LOH	LOH	ROH	ROH	H	H	H	M	N
K-59L	MYCL1, S242		E/L	ROH	H	ROH	H	LOH	LOH	LOH	LOH	ROH	H	H	FS	H	LOH	LOH	N	N
K-61L	MYCL1		E/L	ROH	H	LOH	H	LOH	LOH	H	LOH	FS	LOH	ROH	FS	ROH	H	LOH	M	N
K-78L	MYCL1, S321		E/L	ROH	FS	H	H	ROH	ROH	H	LOH	FS	LOH	LOH	FS	ROH	H	LOH	N	N
K-80L	MYCL1		E/L	H	H	ROH	H	LOH	H	H	ROH	LOH	FS	LOH	LOH	H	H	H	N	N
K-122L	S82		E/L	ROH	FS	ROH	H	ROH	H	H	LOH	ROH	LOH	ROH	LOH	H	LOH	LOH	N	N
K-L35	S82		E/L	ROH	H	H	H	ROH	H	H	H	LOH	ROH	ROH	ROH	LOH	H	H	N	N
K-L38	MYCL1	S250	E/L	FS	FS	FS	H	FS	H	LOH	ROH	ROH	ROH	H	H	ROH	H	H	N	N
K-L36	S242		E/L	LOH	H	LOH	FS	H	ROH	H	LOH	H	LOH	FS	LOH	FS	H	H	N	N
K-L25	S82, S242		E/L	LOH	FS	H	H	H	ROH	LOH	LOH	LOH	FS	LOH	H	LOH	H	H	M	N
K-L9	S321	S123	E/L	H	H	FS	H	FS	LOH	LOH	FS	ROH	LOH	FS	LOH	ROH	H	H	N	M
T-L1	S321		E/L	FS	H	FS	H	H	LOH	ROH	H	FS	FS	ROH	LOH	FS	H	H	N	N
T-L2	S394		E/L	ROH	FS	H	H	ROH	LOH	H	LOH	H	LOH	LOH	LOH	ROH	LOH	LOH	N	N
T-L11	S321		E/L	ROH	H	FS	H	H	H	H	LOH	LOH	LOH	LOH	LOH	H	LOH	ROH	M	N
T-L18		S123, S69	E/L	LOH	H	H	H	H	LOH	H	H	ROH	LOH	LOH	LOH	LOH	H	ROH	N	N
T-L21		S69	E/L	LOH	H	LOH	H	LOH	H	H	LOH	H	LOH	LOH	ROH	LOH	H	H	N	N
T-L23	S242		E/L	LOH	H	LOH	H	FS	H	H	ROH	ROH	LOH	LOH	FS	FS	ROH	H	M	N
T-L27	S321		E/L	H	FS	FS	H	H	ROH	H	LOH	LOH	H	LOH	FS	H	H	H	N	N
K-L2	S82		E/L	FS	LOH	LOH	H	LOH	ROH	LOH	LOH	LOH	H	LOH	LOH	LOH	LOH	ROH	N	N
K-22L	S321		E/L	ROH	H	LOH	H	LOH	ROH	FS	ROH	H	FS	FS	ROH	ROH	ROH	H	M	N
K-63L		S123	E/L	FS	H	H	H	FS	FS	H	H	LOH	LOH	LOH	H	FS	LOH	H	M	N
K-77L	S82		E/L	LOH	FS	FS	H	ROH	ROH	H	H	LOH	LOH	LOH	LOH	ROH	LOH	LOH	N	N
K-121L	S85	S346	E/L	ROH	H	LOH	H	LOH	ROH	H	H	ROH	LOH	ROH	ROH	ROH	LOH	ROH	N	N
K-L34	S82		E/L	LOH	LOH	LOH	H	LOH	ROH	H	LOH	ROH	LOH	LOH	ROH	LOH	H	H	M	N
T-L5	MYCL1		E/L	H	H	FS	H	LOH	LOH	H	ROH	H	LOH	LOH	LOH	LOH	H	H	N	N
T-L15	S85, S321	S69	E/L	LOH	H	H	H	LOH	ROH	LOH	H	ROH	ROH	LOH	LOH	H	H	H	N	N
T-L20	S321		E/L	LOH	FS	H	H	FS	H	LOH	H	ROH	LOH	H	LOH	H	ROH	H	N	N
T-L12	S242	S123, S69	E/L	FS	H	FS	H	H	FS	H	H	ROH	H	H	H	FS	FS	ROH	N	N

FIGURE 1B (CONTINUED).



	EMAST Markers	NCI Markers																	Mutation at	
LMs	FS at	FS at	Genotype	*RBM47*	*WIPF2*	*ANKRD5*	*ZNF161*	*S11*	*S303*	*KDM6B*	*MNT*	*S1179*	*BCL2*	*SEMA6D*	*S818*	*STYK1*	*BCL6B*	*SMARCA2*	*KRAS*	*BRAF*
K-L15			MS	FS	FS	H	H	H	H	H	ROH	LOH	H	H	LOH	H	LOH	LOH	N	N
K-L11			MS	H	ROH	LOH	H	H	LOH	H	LOH	FS	LOH	H	LOH	LOH	H	H	N	N
T-L16			MS	FS	H	ROH	H	ROH	H	LOH	LOH	ROH	LOH	H	LOH	LOH	H	ROH	N	N
K-L19			MS	LOH	LOH	LOH	H	ROH	ROH	H	LOH	H	LOH	LOH	ROH	H	H	H	N	N
K-L26			MS	ROH	H	ROH	H	LOH	ROH	H	ROH	ROH	LOH	H	LOH	H	LOH	ROH	N	N
K-L33			MS	LOH	H	H	H	LOH	ROH	H	LOH	H	LOH	LOH	LOH	ROH	LOH	H	M	N
K-L23			MS	LOH	H	H	H	H	ROH	H	LOH	LOH	LOH	ROH	H	ROH	H	ROH	N	N
K-L30			MS	LOH	H	LOH	H	ROH	LOH	H	LOH	ROH	ROH	ROH	LOH	ROH	H	ROH	M	N
T-L13			MS	H	H	ROH	H	ROH	LOH	H	LOH	ROH	ROH	ROH	ROH	H	ROH	ROH	M	N
T-L22			MS	ROH	H	LOH	H	H	LOH	H	H	H	LOH	LOH	LOH	H	LOH	H	M	N
T-L25			MS	LOH	LOH	H	H	H	H	LOH	LOH	LOH	LOH	ROH	LOH	ROH	H	H	N	N
T-L32			MS	LOH	H	ROH	H	LOH	ROH	H	H	ROH	LOH	H	LOH	ROH	H	ROH	N	N
K-L10			MS	ROH	H	LOH	H	H	ROH	H	ROH	ROH	H	ROH	ROH	ROH	H	ROH	N	N
K-L3			MS	FS	H	H	H	H	LOH	H	LOH	ROH	H	ROH	LOH	H	H	ROH	M	N
K-L39			MS	LOH	H	ROH	H	LOH	ROH	H	LOH	LOH	LOH	LOH	LOH	ROH	LOH	H	N	N
K-L31			MS	LOH	LOH	H	H	LOH	LOH	LOH	LOH	LOH	LOH	LOH	ROH	H	H	ROH	N	N
T-L9			MS	LOH	H	H	H	LOH	LOH	LOH	H	ROH	ROH	H	LOH	ROH	H	ROH	M	M
T-L31			MS	H	H	LOH	H	LOH	ROH	LOH	LOH	LOH	LOH	LOH	LOH	LOH	H	H	N	N
K-L4			MS	H	H	H	H	LOH	ROH	H	H	LOH	ROH	LOH	ROH	H	H	ROH	M	N
K-L22			MS	H	H	ROH	H	LOH	ROH	LOH	ROH	LOH	H	ROH	ROH	H	H	ROH	N	N
K-44L			MS	LOH	H	LOH	H	LOH	ROH	H	LOH	ROH	H	LOH	LOH	ROH	H	H	N	N
T-L26			MS	ROH	H	LOH	H	LOH	ROH	ROH	ROH	H	ROH	ROH	LOH	LOH	H	H	M	N

MSI-H LM (1)

T-L10	S85, S321	S123, BAT25 BAT26	MSI-H	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	M	N

Partially characterized LMs (2)

K-L12	S82		E/L	N.D.	ROH	H	N.D.	ROH	FS	N.D.	LOH	N.D.	LOH	ROH	N.D.	N.D.	LOH	N.D.	N	N
K-L20		S250	E/L	N.D.	N.D.	N.D.	H	N.D.	N.D.	LOH	N.D.	N.D.	N.D.	N.D.	LOH	N.D.	N.D.	ROH	N.D.	N.D.
T-L8	S85, S82, S242	S69	E/L	LOH	N.D.	LOH	N.D.	ROH	N.D.	N.D.	LOH	N.D.	LOH	LOH	N.D.	ROH	H	N.D.	N.D.	N.D.

SUPPLEMENTARY MATERIALS: FIGURES

Supplementary Fig. 1

A: KRAS/BRAF mutations and Frame-shift (FS) and/or LOH of selected markers in 24 E/L positive LMs from primary CRC

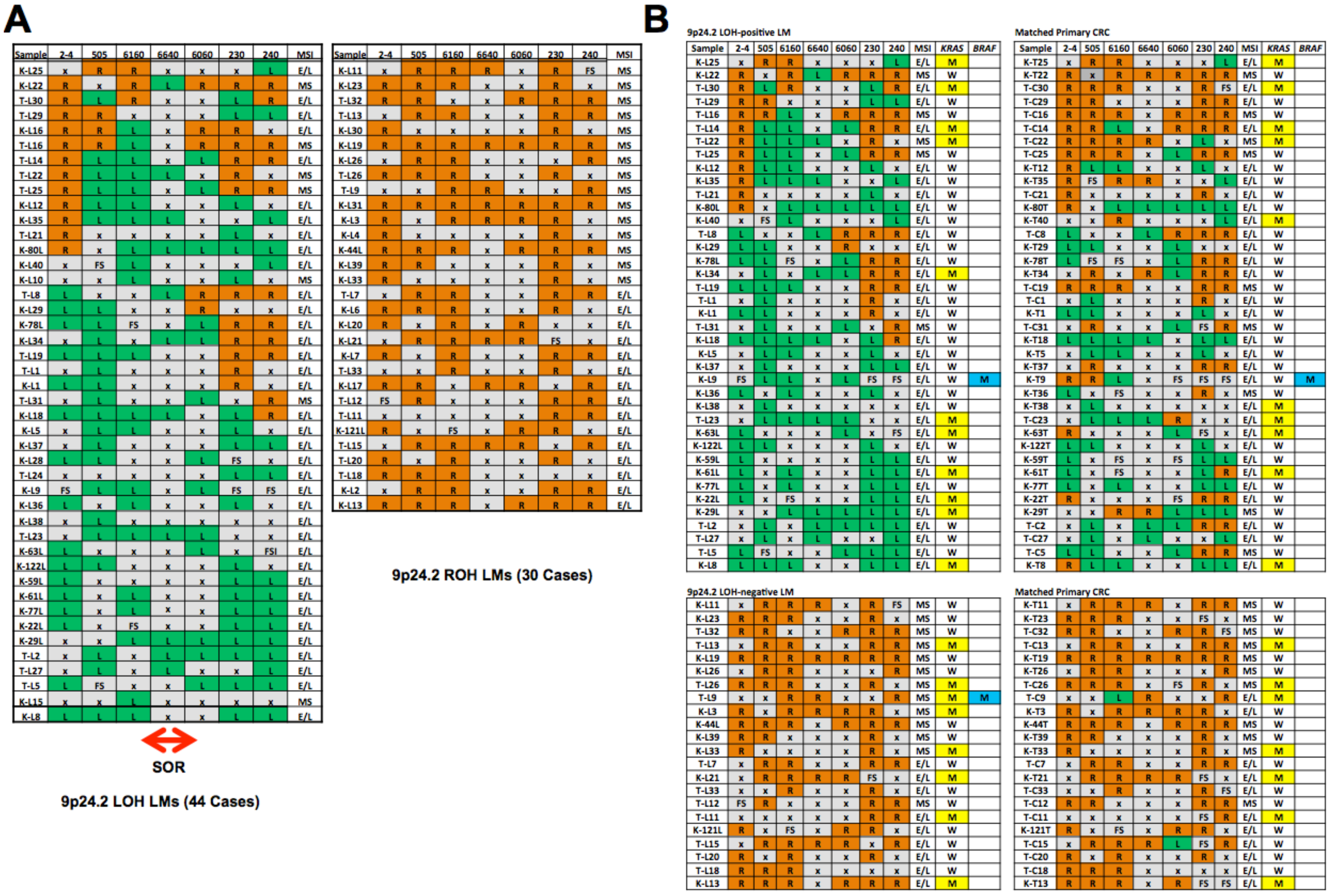
The columns depict the following: Sample ID, FS status at EMAST markers, FS status at NCI markers, Genotype, FS and/or LOH data of 15 loci and mutation data for KRAS/BRAF. For Sample ID, the samples with K number and with T number were collected at DPCNU (Kwanju, Korea) and at DPOHTU (Tokyo, Japan) respectively. For FS status at EMAST and NCI markers, the markers showing FS are listed. For genotype, all 24 LMs were categorized as E/L according to our definition in **Materials and Methods**. For FS and/or LOH data of 15 loci, a brown box with ROH indicates retention of heterozygosity; an empty color box with H indicates cases that were not informative; a yellow box with FS indicates cases with frame-shift, and a green box with LOH indicates loss of heterozygosity. Following abbreviations were used for markers *D2S123* (S123), *D17S250* (S250), *D18S64* (S64), and *D18S69* (S69), *D20S82* (S82), *D20S85* (S85)*, D8S321* (S321), *D9S242* (S242), *D19S394* (S394), *D21S11* (S11), *D9S303* (S303), *D5S818* (S818) and *D8S1179* (S1179). For KRAS/BRAF mutations, a blue box with M represents presence of mutations whereas an empty color box with N represents no evidence of mutations.

B: FS and LOH status of selected markers in additional 55 LMs from primary CRC

The columns depict the same as the columns in Fig. S1A. . For genotype, 55 LMs were categorized as E/L, MS or MSI-H according to our definition in **Materials and Methods**. For FS and/or LOH data of 15 loci, an empty color box with N.D. indicates the status was not determined.

(1): T-L10 exhibited MSI-H genotype. Therefore, MSI or LOH at 15 loci was not determined.

(2): Template DNA from K-L12, L-L20 and T-L8 was not enough to analyze all of 15 loci.



Sample	2-4	505	6160	6640	6060	230	240	MSI
K-L25	x	R	R	x	x	x	L	E/L
K-L22	R	x	R	L	R	R	R	MS
T-L30	R	L	R	x	x	L	R	E/L
T-L29	R	R	x	x	x	L	L	E/L
K-L16	R	R	L	x	R	R	x	E/L
T-L16	R	R	L	x	R	R	R	MS
T-L14	R	L	L	x	L	R	R	E/L
T-L22	R	L	L	L	x	R	x	MS
T-L25	R	L	L	x	L	R	R	MS
K-L12	R	L	L	x	x	L	x	E/L
K-L35	R	L	L	L	x	x	L	E/L
T-L21	R	x	x	x	x	L	x	E/L
K-80L	R	x	L	L	L	L	L	E/L
K-L40	x	FS	L	x	x	x	L	E/L
K-L10	x	x	L	x	x	L	x	MS
T-L8	L	x	x	L	R	R	R	E/L
K-L29	L	L	x	x	R	x	x	E/L
K-78L	L	L	FS	x	L	R	R	E/L
K-L34	x	L	x	L	L	R	R	E/L
T-L19	L	L	L	x	x	R	R	E/L
T-L1	x	L	x	x	x	R	x	E/L
K-L1	L	L	x	x	x	R	x	E/L
T-L31	x	L	x	x	L	x	R	MS
K-L18	L	L	L	L	x	L	R	E/L
K-L5	x	L	L	x	x	L	x	E/L
K-L37	x	L	x	x	x	L	L	E/L
K-L28	L	L	x	x	L	FS	x	E/L
T-L24	x	x	x	x	x	L	L	E/L
K-L9	FS	L	L	x	L	FS	FS	E/L
K-L36	L	x	L	x	x	L	x	E/L
K-L38	x	L	x	x	x	x	x	E/L
T-L23	x	L	L	L	L	x	x	E/L
K-63L	L	x	x	x	L	x	FSI	E/L
K-122L	L	L	x	x	x	L	x	E/L
K-59L	L	x	x	x	x	L	L	E/L
K-61L	L	x	L	x	x	L	L	E/L
K-77L	L	x	L	x	x	L	L	E/L
K-22L	L	x	FS	x	x	L	L	E/L
K-29L	x	x	L	L	L	L	L	E/L
T-L2	x	L	x	L	L	L	L	E/L
T-L27	x	L	x	L	x	x	L	E/L
T-L5	L	FS	x	x	L	L	L	E/L
K-L15	x	x	L	x	x	x	x	MS
K-L8	L	L	L	x	x	L	L	E/L

9p24.2 LOH LMs (44 Cases)

Sample	2-4	505	6160	6640	6060	230	240	MSI
K-L11	x	R	R	R	x	R	FS	MS
K-L23	R	R	R	x	x	R	x	MS
T-L32	R	R	x	x	R	R	R	MS
T-L13	x	R	R	x	x	R	R	MS
K-L30	R	x	x	x	x	R	x	MS
K-L19	R	R	R	R	R	R	R	MS
K-L26	x	R	R	x	x	x	R	MS
T-L26	R	R	R	x	x	R	x	MS
T-L9	x	x	R	R	x	x	R	MS
K-L31	R	R	R	R	R	R	R	MS
K-L3	R	x	R	R	R	R	x	MS
K-L4	R	x	x	x	x	R	x	MS
K-44L	R	R	R	x	R	R	R	MS
K-L39	R	R	x	x	x	R	x	MS
K-L33	R	x	x	x	x	R	x	MS
T-L7	x	R	R	x	x	R	R	E/L
K-L6	R	R	R	x	x	R	x	E/L
K-L20	R	x	R	x	R	x	x	E/L
K-L21	x	R	R	R	R	FS	x	E/L
K-L7	R	x	R	x	x	R	R	E/L
T-L33	x	x	R	x	x	R	x	E/L
K-L17	R	R	x	R	R	x	R	E/L
T-L12	FS	R	x	x	x	R	R	E/L
T-L11	x	x	x	x	x	R	R	E/L
K-121L	R	x	FS	x	R	R	x	E/L
T-L15	x	R	R	R	R	x	R	E/L
T-L20	R	x	R	x	x	R	x	E/L
T-L18	R	R	R	x	x	x	x	E/L
K-L2	x	R	R	x	x	R	R	E/L
K-L13	R	R	R	x	R	R	R	E/L

9p24.2 ROH LMs (30 Cases)

9p24.2 LOH-positive LM

Sample	2-4	505	6160	6640	6060	230	240	MSI	*KRAS*	*BRAF*
K-L25	x	R	R	x	x	x	L	E/L	M	
K-L22	R	x	R	L	R	R	R	MS	W	
T-L30	R	L	R	x	x	L	R	E/L	M	
T-L29	R	R	x	x	x	L	L	E/L	W	
T-L16	R	R	L	x	R	R	R	MS	W	
T-L14	R	L	L	x	L	R	R	E/L	M	
T-L22	R	L	L	L	x	R	x	MS	M	
T-L25	R	L	L	x	L	R	R	MS	W	
K-L12	R	L	L	x	x	L	x	E/L	W	
K-L35	R	L	L	L	x	x	L	E/L	W	
T-L21	R	x	x	x	x	L	x	E/L	W	
K-80L	R	x	L	L	L	L	L	E/L	W	
K-L40	x	FS	L	x	x	x	L	E/L	W	
T-L8	L	x	x	L	R	R	R	E/L	W	
K-L29	L	L	x	x	R	x	x	E/L	W	
K-78L	L	L	FS	x	L	R	R	E/L	W	
K-L34	x	L	x	L	L	R	R	E/L	M	
T-L19	L	L	L	x	x	R	R	E/L	W	
T-L1	x	L	x	x	x	R	x	E/L	W	
K-L1	L	L	x	x	x	R	x	E/L	W	
T-L31	x	L	x	x	L	x	R	MS	W	
K-L18	L	L	L	L	x	L	R	E/L	W	
K-L5	x	L	L	x	x	L	x	E/L	W	
K-L37	x	L	x	x	x	L	L	E/L	W	
K-L9	FS	L	L	x	L	FS	FS	E/L	W	M
K-L36	L	x	L	x	x	L	x	E/L	W	
K-L38	x	L	x	x	x	x	x	E/L	W	
T-L23	x	L	L	L	L	x	x	E/L	M	
K-63L	L	x	x	x	L	x	FS	E/L	M	
K-122L	L	L	x	x	x	L	x	E/L	W	
K-59L	L	x	x	x	x	L	L	E/L	W	
K-61L	L	x	L	x	x	L	L	E/L	M	
K-77L	L	x	L	x	x	L	L	E/L	W	
K-22L	L	x	FS	x	x	L	L	E/L	M	
K-29L	x	x	L	L	L	L	L	E/L	M	
T-L2	x	L	x	L	L	L	L	E/L	W	
T-L27	x	L	x	L	x	x	L	E/L	W	
T-L5	L	FS	x	x	L	L	L	E/L	W	
K-L8	L	L	L	x	x	L	L	E/L	M	

Matched Primary CRC

Sample	2-4	505	6160	6640	6060	230	240	MSI	*KRAS*	*BRAF*
K-T25	x	R	R	x	x	x	L	E/L	M	
K-T22	R	x	R	R	R	R	R	MS	W	
T-C30	R	R	R	x	x	R	FS	E/L	M	
T-C29	R	R	x	x	x	R	R	E/L	W	
T-C16	R	R	R	x	R	R	R	MS	W	
T-C14	R	R	L	x	R	R	R	E/L	M	
T-C22	R	R	R	R	x	L	x	MS	M	
T-C25	R	R	R	x	L	R	R	MS	W	
K-T12	R	L	L	x	x	L	x	E/L	W	
K-T35	R	FS	R	R	x	x	L	E/L	W	
T-C21	R	x	x	x	x	R	x	E/L	W	
K-80T	R	x	L	L	L	L	L	E/L	W	
K-T40	x	x	R	x	x	x	L	E/L	M	
T-C8	L	x	x	L	R	R	R	E/L	W	
K-T29	L	L	x	x	L	x	x	E/L	W	
K-78T	L	FS	FS	x	L	R	R	E/L	W	
K-T34	x	R	x	R	L	R	R	E/L	W	
T-C19	R	R	R	x	x	R	R	MS	W	
T-C1	x	L	x	x	x	R	x	E/L	W	
K-T1	L	L	x	x	x	L	x	E/L	W	
T-C31	x	R	x	x	L	FS	R	MS	W	
K-T18	L	L	L	L	x	L	L	E/L	W	
K-T5	x	L	L	x	x	L	x	E/L	W	
K-T37	x	R	x	x	x	R	R	E/L	W	
K-T9	R	R	L	x	FS	FS	FS	E/L	W	M
K-T36	L	x	FS	x	x	R	x	MS	W	
K-T38	x	L	x	x	x	x	x	E/L	M	
T-C23	x	L	L	L	R	x	x	E/L	M	
K-63T	R	x	x	x	L	FS	x	E/L	M	
K-122T	L	L	x	x	x	L	x	E/L	W	
K-59T	L	x	FS	x	FS	L	L	E/L	W	
K-61T	L	x	FS	x	x	L	R	E/L	M	
K-77T	L	x	L	x	x	L	L	E/L	W	
K-22T	R	x	x	x	FS	R	R	E/L	W	
K-29T	x	x	R	R	L	L	L	MS	W	
T-C2	x	L	x	L	L	R	R	E/L	W	
T-C27	x	L	x	L	x	x	L	E/L	W	
T-C5	L	L	x	x	L	R	R	MS	W	
K-T8	R	L	L	x	x	L	L	E/L	M	

9p24.2 LOH-negative LM

K-L11	x	R	R	R	x	R	FS	MS	W	
K-L23	R	R	R	x	x	R	x	MS	W	
T-L32	R	R	x	x	R	R	R	MS	W	
T-L13	x	R	R	x	x	R	R	MS	M	
K-L19	R	R	R	R	R	R	R	MS	W	
K-L26	x	R	R	x	x	x	R	MS	W	
T-L26	R	R	R	x	x	R	x	MS	M	
T-L9	x	x	R	R	x	x	R	MS	M	M
K-L3	R	x	R	R	R	R	x	MS	M	
K-44L	R	R	R	x	R	R	R	MS	W	
K-L39	R	R	x	x	x	R	x	MS	W	
K-L33	R	x	x	x	x	R	x	MS	M	
T-L7	x	R	R	x	x	R	R	E/L	W	
K-L21	x	R	R	R	R	FS	x	E/L	M	
T-L33	x	x	R	x	x	R	x	E/L	W	
T-L12	FS	R	x	x	x	R	R	MS	W	
T-L11	x	x	x	x	x	R	R	E/L	M	
K-121L	R	x	FS	x	R	R	x	E/L	W	
T-L15	x	R	R	R	R	x	R	E/L	W	
T-L20	R	x	R	x	x	R	x	E/L	W	
T-L18	R	R	R	x	x	x	x	E/L	W	
K-L13	R	R	R	x	R	R	R	E/L	M	

Matched Primary CRC

K-T11	x	R	R	R	x	R	R	MS	W	
K-T23	R	R	R	x	x	FS	x	MS	W	
T-C32	R	R	x	x	R	R	FS	MS	W	
T-C13	x	R	R	x	x	R	R	MS	M	
K-T19	R	R	R	R	R	R	R	MS	W	
K-T26	x	R	R	x	x	x	R	MS	W	
T-C26	R	R	R	x	FS	R	x	MS	M	
T-C9	x	x	L	R	x	x	R	E/L	M	
K-T3	R	x	R	R	R	R	x	E/L	W	
K-44T	R	R	R	x	R	R	R	MS	W	
K-T39	R	R	x	x	x	R	x	MS	W	
K-T33	R	x	x	x	x	R	x	MS	M	
T-C7	x	R	R	x	x	R	R	E/L	W	
K-T21	x	R	R	R	R	FS	x	E/L	M	
T-C33	x	x	R	x	x	R	FS	E/L	W	
T-C12	R	R	x	x	x	R	R	MS	W	
T-C11	x	x	x	x	x	FS	R	E/L	M	
K-121T	R	x	FS	x	R	R	x	E/L	W	
T-C15	x	R	R	R	L	FS	R	E/L	W	
T-C20	R	x	R	x	x	R	x	E/L	W	
T-C18	R	R	R	x	x	x	x	E/L	W	
K-T13	R	R	R	x	R	FS	FS	E/L	M	

Supplementary Fig. 2

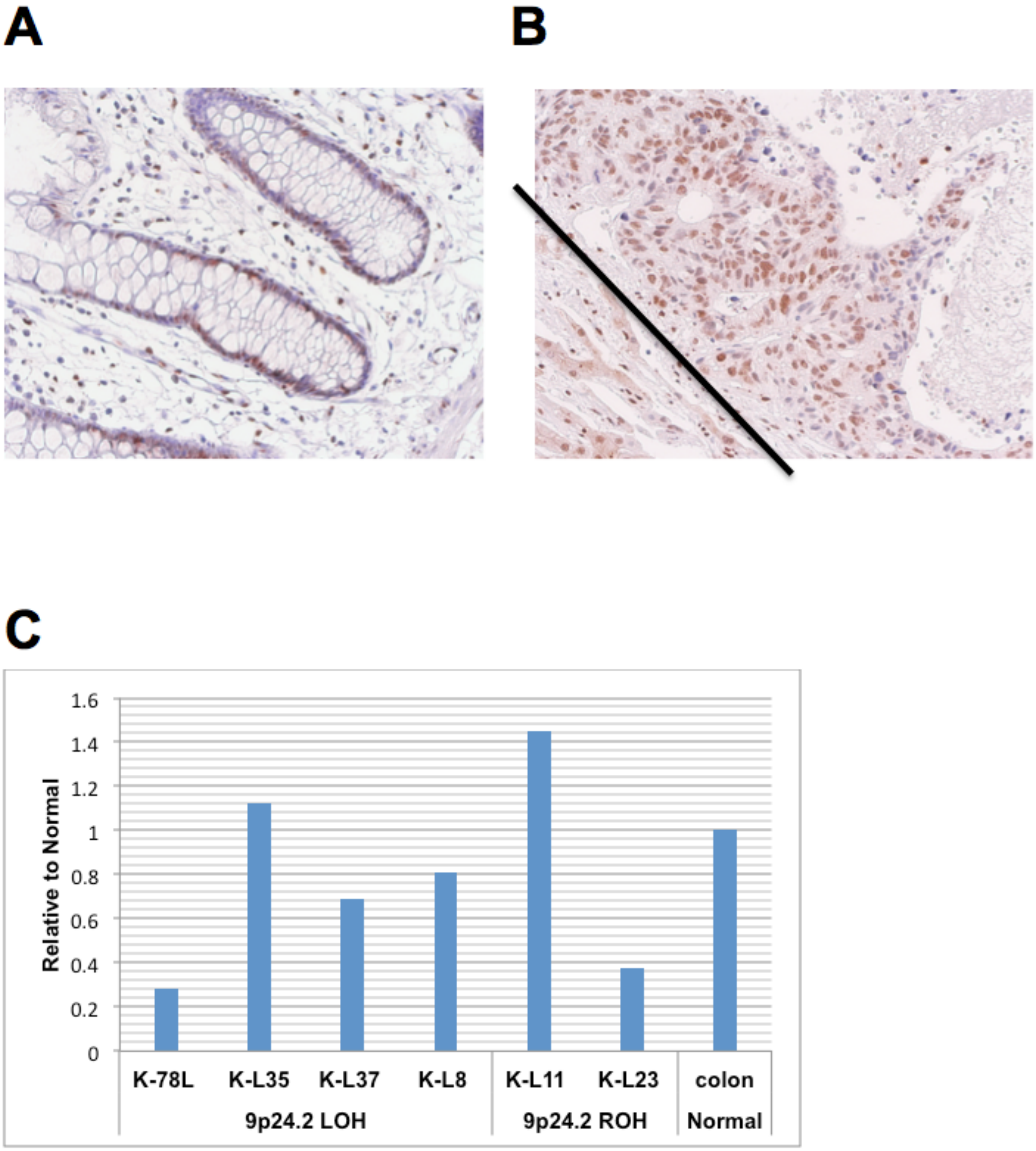
A: FS and 9p24.2 LOH status of 74 LMs

This figure provides detailed data from 74 LMs. The columns depict the following: LOH data for 7 markers (*2-4, 505, 6160, 6640, 6060, 230*, and *240*) and MSI status. For LOH data, a brown box with **R** indicates retention of heterozygosity (ROH); a gray box with **X** indicates cases that were not informative; a gray box with FS indicates cases with frame-shifts, and a green box with **L** indicates loss of heterozygosity (LOH). For MSI status **E/L** indicates EMAST/MSI-L and MS indicates cases with highly stable microsatellites. Two markers, *2-4* and *505*, localized within the *SMARCA2* gene. The smallest overlapping region of LOH among 9p24.2 LOH-positive LMs is banded by *6160* and *6060* (red arrow).

B: FS, 9p24.2 LOH, and *KRAS/BRAF* mutation status in 61 matching LMs and metastatic primary CRCs

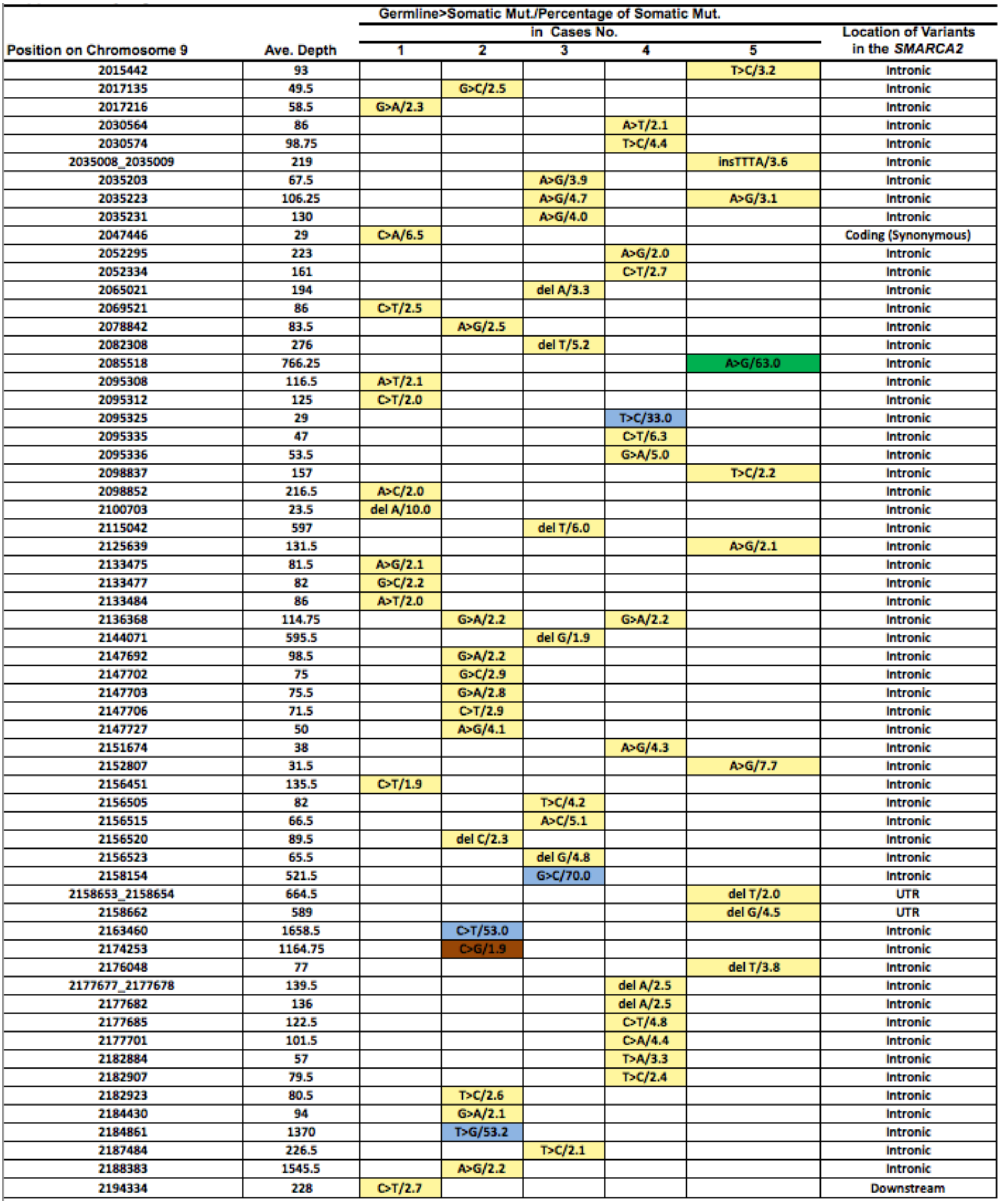
This figure provides detailed data from 61 LMs (left column) and matching primary CRCs (right column). In contrast to the LMs, no SOR was detected in matching primary CRC.

SUPPLEMENTARY MATERIALS: FIGURES



Supplementary Fig. 3 Expression of SMARCA2 in normal human colon and LMs from primary CRC Formalin-fixed paraffin-embedded (FFPE) tissues were deparaffinized and rehydrated. After antigen retrieval [121°C for 15 min in 0.01 mol/L citrate buffer (pH 6.0)], the tissues were treated with anti-SMARCA2 antibody (1:250, HPA029981, Sigma) overnight at 4°C, followed by incubation with secondary antibody. Staining was developed by diaminobenzide chromogen and counterstained with hematoxylin. The brown signals in nucleus indicated the presence of SMARCA2. Normal colon tissues and six 9p24.2 LOH-positive LM tissues were examined. There was no difference in the levels of SMARCA2 expression between normal colon and 9p24.2 LOH-positive LMs. **A:** Normal colon, **B:** 9p24.2 LOH-positive LM. A metastasis tissue locates at upper right side of a black line and normal liver cells are visible at lower left side of a black line. **C:** mRNAs were isolated from FFPE LMs and normal tissues and were subjected to cDNA synthesis. qRT-PCR was performed to compare the expression levels of *SMARCA2* in LMs exhibiting 9p24.2 LOH (K-78L, K-L35, K-L37 and K-L8) or 9p24.2 ROH (K-L11 and k-L23) to those in normal colon tissues. For normalization, the expressions of house-keeping genes *TXNR1* and *PSBM6* were used. Relative expression of the *SMARCA2* in LM vs normal colon was determined by qBase^{+} software (Biogazelle, Belgium). No correlation between the levels of *SMARCA2* mRNA and presence or absence of 9p24.2 LOH in LMS was observed.

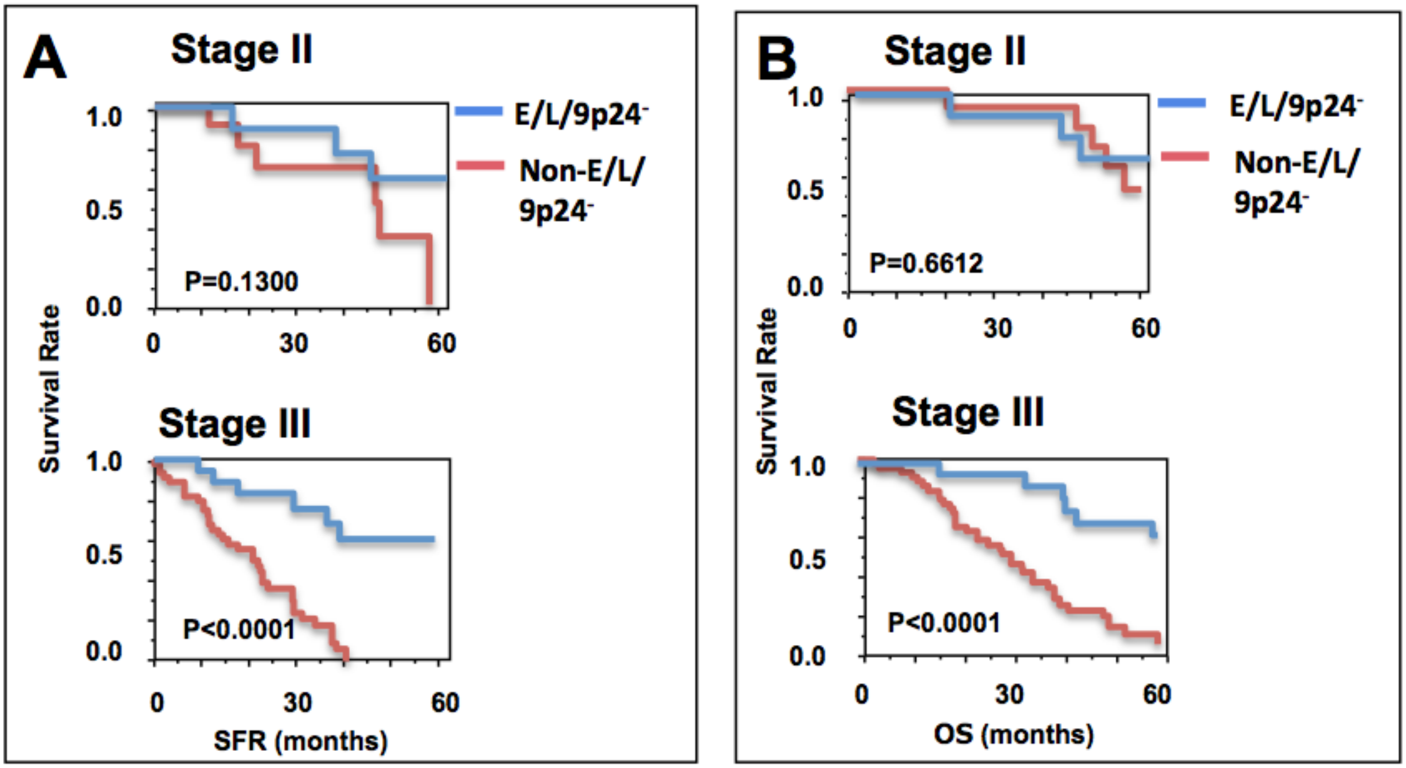
SUPPLEMENTARY MATERIALS: FIGURES



Position on Chromosome 9	Ave. Depth	Germline>Somatic Mut./Percentage of Somatic Mut. in Cases No.					Location of Variants in the *SMARCA2*
		1	2	3	4	5	
2015442	93					T>C/3.2	Intronic
2017135	49.5		G>C/2.5				Intronic
2017216	58.5	G>A/2.3					Intronic
2030564	86				A>T/2.1		Intronic
2030574	98.75				T>C/4.4		Intronic
2035008_2035009	219					insTTTA/3.6	Intronic
2035203	67.5			A>G/3.9			Intronic
2035223	106.25			A>G/4.7		A>G/3.1	Intronic
2035231	130			A>G/4.0			Intronic
2047446	29	C>A/6.5					Coding (Synonymous)
2052295	223				A>G/2.0		Intronic
2052334	161				C>T/2.7		Intronic
2065021	194			del A/3.3			Intronic
2069521	86	C>T/2.5					Intronic
2078842	83.5		A>G/2.5				Intronic
2082308	276			del T/5.2			Intronic
2085518	766.25					A>G/63.0	Intronic
2095308	116.5	A>T/2.1					Intronic
2095312	125	C>T/2.0					Intronic
2095325	29				T>C/33.0		Intronic
2095335	47				C>T/6.3		Intronic
2095336	53.5				G>A/5.0		Intronic
2098837	157					T>C/2.2	Intronic
2098852	216.5	A>C/2.0					Intronic
2100703	23.5	del A/10.0					Intronic
2115042	597			del T/6.0			Intronic
2125639	131.5					A>G/2.1	Intronic
2133475	81.5	A>G/2.1					Intronic
2133477	82	G>C/2.2					Intronic
2133484	86	A>T/2.0					Intronic
2136368	114.75		G>A/2.2		G>A/2.2		Intronic
2144071	595.5			del G/1.9			Intronic
2147692	98.5		G>A/2.2				Intronic
2147702	75		G>C/2.9				Intronic
2147703	75.5		G>A/2.8				Intronic
2147706	71.5		C>T/2.9				Intronic
2147727	50		A>G/4.1				Intronic
2151674	38				A>G/4.3		Intronic
2152807	31.5					A>G/7.7	Intronic
2156451	135.5	C>T/1.9					Intronic
2156505	82			T>C/4.2			Intronic
2156515	66.5			A>C/5.1			Intronic
2156520	89.5		del C/2.3				Intronic
2156523	65.5			del G/4.8			Intronic
2158154	521.5			G>C/70.0			Intronic
2158653_2158654	664.5					del T/2.0	UTR
2158662	589					del G/4.5	UTR
2163460	1658.5		C>T/53.0				Intronic
2174253	1164.75		C>G/1.9				Intronic
2176048	77					del T/3.8	Intronic
2177677_2177678	139.5				del A/2.5		Intronic
2177682	136				del A/2.5		Intronic
2177685	122.5				C>T/4.8		Intronic
2177701	101.5				C>A/4.4		Intronic
2182884	57				T>A/3.3		Intronic
2182907	79.5				T>C/2.4		Intronic
2182923	80.5		T>C/2.6				Intronic
2184430	94		G>A/2.1				Intronic
2184861	1370		T>G/53.2				Intronic
2187484	226.5			T>C/2.1			Intronic
2188383	1545.5		A>G/2.2				Intronic
2194334	228	C>T/2.7					Downstream

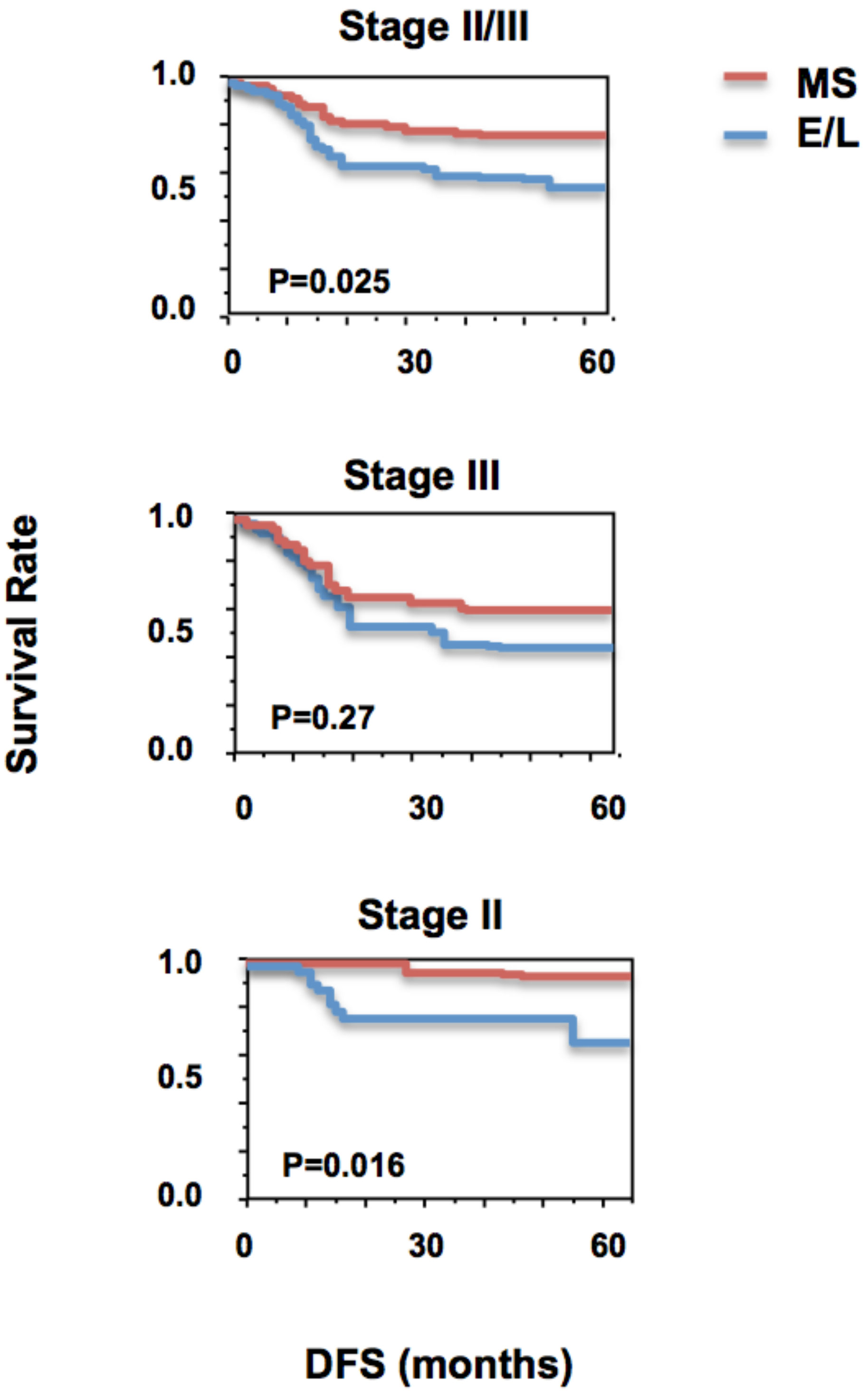
Supplementary Fig. 4

Targeted deep sequencing of the *SMARCA2* locus To determine whether tumor-specific recurrent sequence variants within the *SMARCA2* region exist, genomic DNAs isolated from five matching normal/tumor pairs exhibiting 9p24.2 LOH were subjected to a targeted deep sequencing. A region of about 180 kb containing the *SMARCA2* locus (Chr9: 2,015,311-2,194,334) was analyzed after sequencing by Hybrid Capture method with the SureSelect Target Enrichment System and on Illuminainstruments. Variant calling was performed by Mutect and VarScan. The columns depict the following: position of a sequence variant on chromosome 9, average depth of sequence reading, sequence variants and their frequencies in each sample, and location of variant within the *SMARCA2* gene. For variants and their frequencies, a variant in a green box is called by both Mutect and VarScan, one in a blue box is called by VarScan with high confidence, a variant in a yellow box is called by VarScan only with low confidence, and the one in brown box is called by Mutect only. For a variant location within *SMARCA2*, an each variant was located at intronic, coding, 3' untranslated region (UTR), or down-stream region. Although 64 sequence variants were detected among 5 tumor samples when compared to matching normal samples, most of them were found in intron of the *SMARCA2* and at a low frequency (1.9%~30%) suggesting that these alterations might have occurred after LOH event. Even though 4 sequence variants with high frequency (more than 50%) were detected, they are at the intron and sporadic among 5 tumor samples. We conclude that there was no recurrent mutation in the *SMARCA2* locus in 9p24.2 LOH-positive tumors.



Supplementary Fig.5

Kaplan-Meier curves for SFR or OS of patients with Stage III (**n=67**) or patients with Stage II (**n=21**). The SFR was compared between patients with primary CRC exhibiting E/L/9p24$^-$ (blue line: Stage III, n=45; Stage II, n=9) and non-E/L/p29$^-$ (red line: Stage III, n=22; Stage II, n=12) in Stage II (**upper**) or III (**bottom)** CRC. (**A**), The OS was compared between patients with primary CRC exhibiting E/L/9p24- (blue line) and non-E/L/p29- (red line) in Stage II (**upper**) or III (**bottom)** (**B).** The X-axis represents SFR or OS in months. The Y-axis represents survival rates. P values were determined by log-rank test.



Supplementary Fig. 6
Kaplan-Meier curves for DFS of patients with Stage II/III (n=124), Stage III (n=75) and Stage II CRCs (n=50). DFS was compared between patients with primary CRC exhibiting E/L (blue line) and non-E/L (red line) in Stage II plus III primary CRC (**upper**), Sage III primary CRC (**middle**) or in Stage II primary CRC (**bottom).** Stage II/III :E/L (n=65) vs non-E/L (n=59), Stage III: E/L (n=41) vs non-E/L (n=34), Stage II: E/L (n=24) vs non-E/L (n=26). The X-axis represents DFS in months. The Y-axis represents survival rates. P values were determined by log-rank test.

SUPPLEMENTARY MATERIALS: FIGURES

Supplementary References

1. Haugen A *et al.* (2008) Genetic instability caused by loss of MutS homologue 3 in human colorectal cancer. *Cancer Res* **68:**8465–8472.
2. Schuelke M. (2000) An economic method for the fluorescnt labeling of PCR fragments. Nat Biotech 18:233-234.
3. Kozlowski P, de Mezer M, Krzyzosiak WJ (2010) Trinucleotide repeats in human genome and exome. Nucleic Acids Res 38:4027-4039.
4. Missirlis PI et al. (2005) Satellog: a database for the identification and prioritization of satellite repeats in disease association studies. BMC Bioinformatics 6:145.
5. Bacolla A et al. (2008) Abundance and length of simple repeats in vertebrate genomes are determined by their structural properties. Genome Res 18:1545-1553.