Supplemental Information for

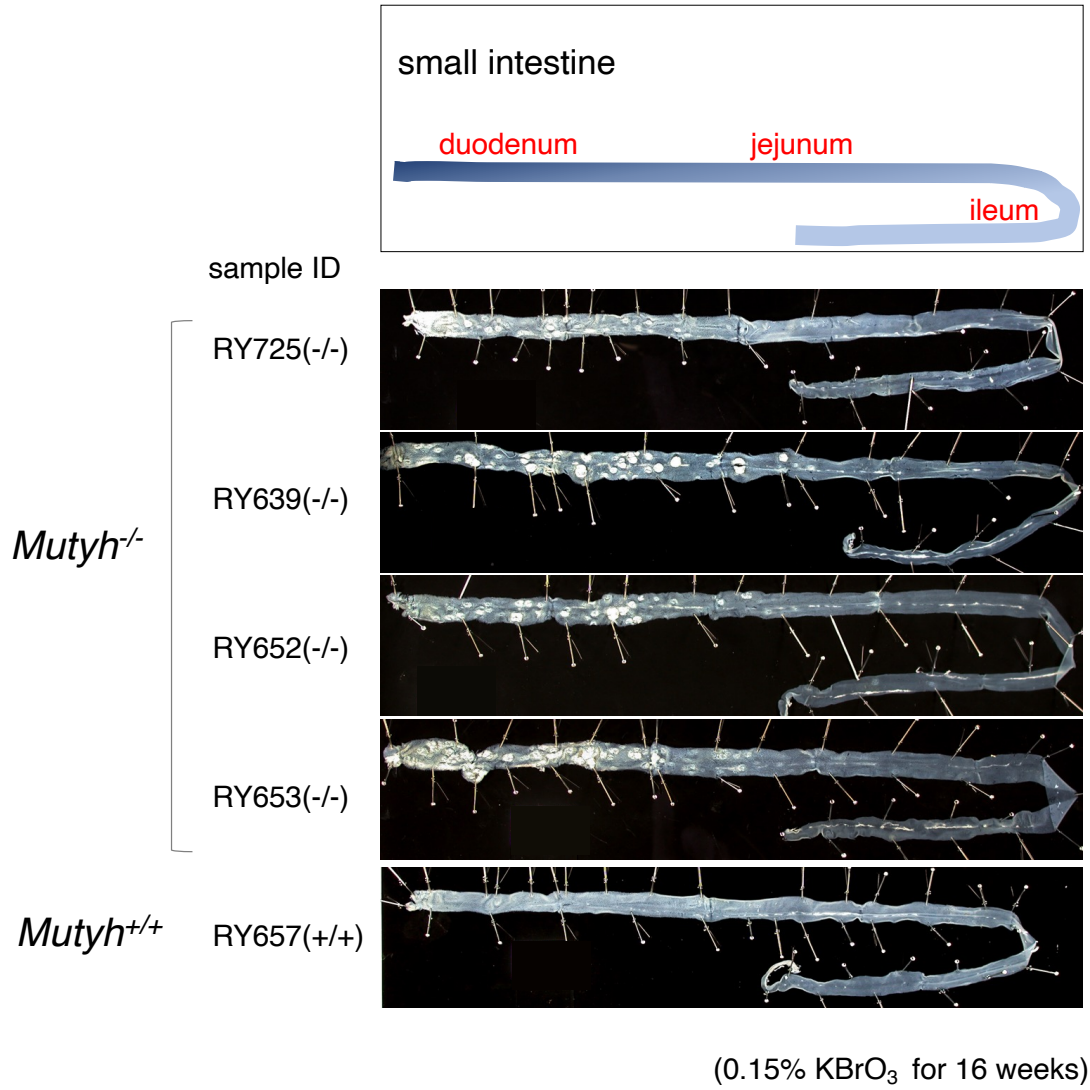
Oxidative stress accelerates intestinal tumorigenesis by enhancing 8-oxoguanine-mediated mutagenesis in MUTYH-deficient mice

Mizuki Ohno*, Noriko Takano, Kyoko Hidaka, Fumiko Sasaki, Kazumi Yamauchi, Yasunobu Aoki, Takehiko Nohmi, Ysaku Nakabeppu, Yoshimichi Nakatsu, Teruhisa Tsuzuki

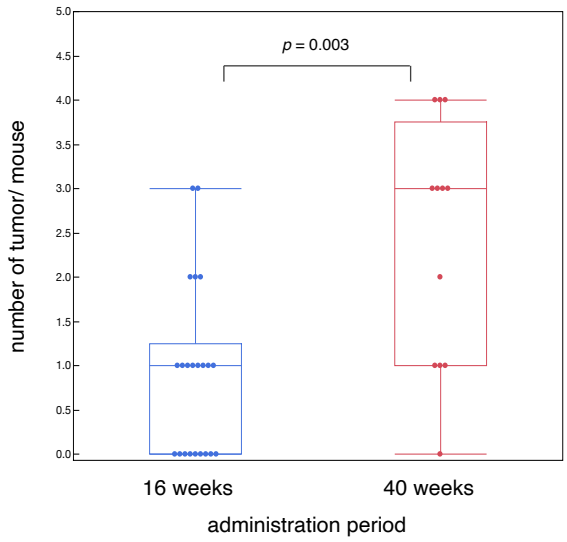
* Correspondence: mizuki.ohno.700@m.kyushu-u.ac.jp

This PDF file includes:

- Supplemental Figures S1 to S11

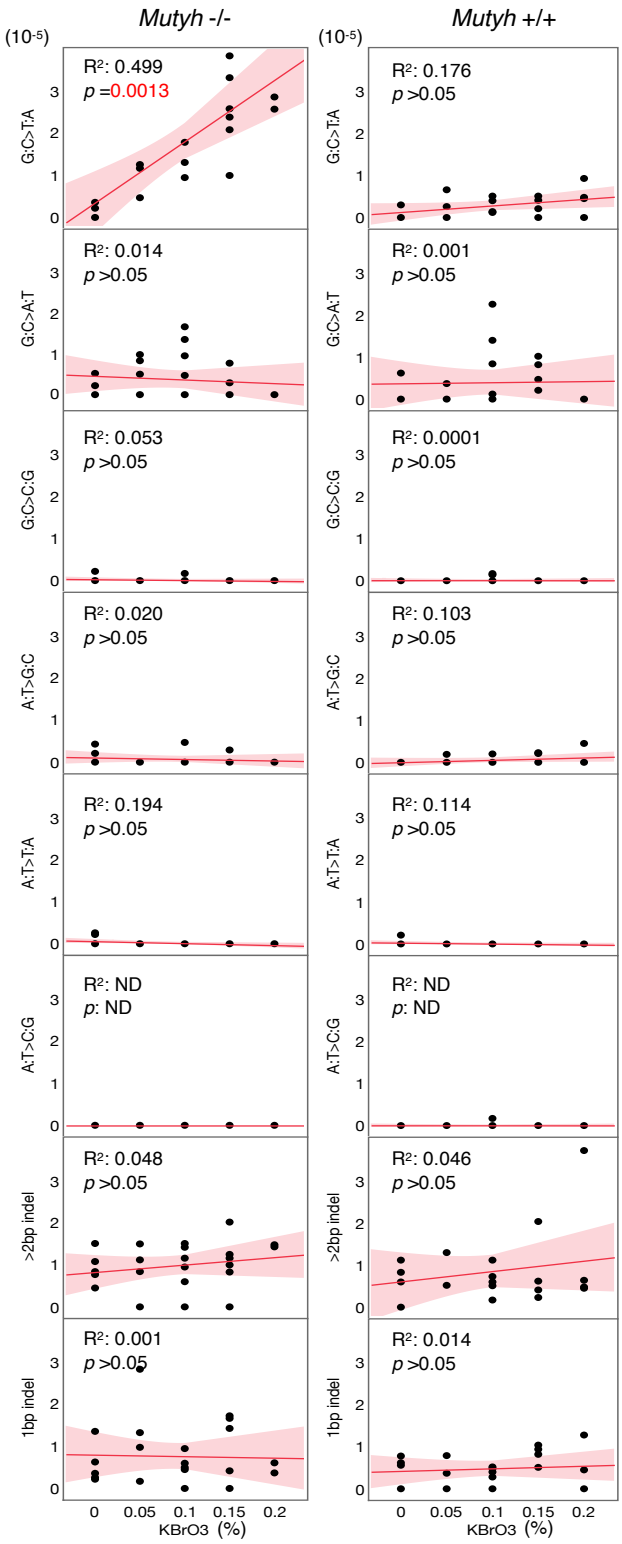


Supplemental Figure S1. Low-magnification image of the small intestines of $KBrO_3$-treated *Mutyh*$^{-/-}$ and *Mutyh*$^{+/+}$ mice. Multiple tumors (white spots in the photos) were observed from the duodenum to jejunum in *Mutyh*$^{-/-}$ mice.

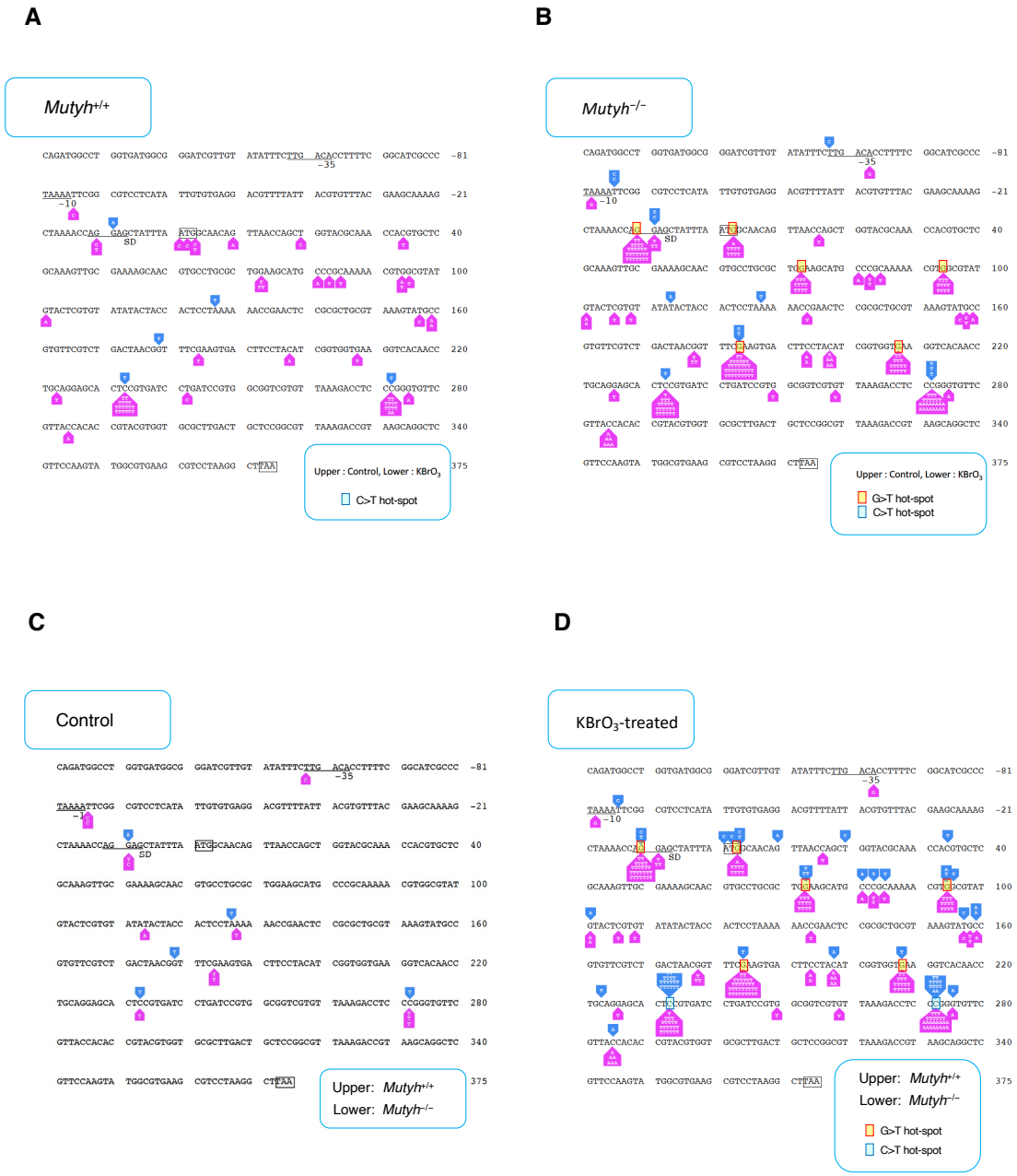


0.15% $KBrO_3$-administration period	number of mice examined	number of tumor /mouse (mean ± SD)	number of mice with tumor(s)	% of mice with tumor(s) (tumor prevalence)
16 weeks	22	0.90 ± 0.97	13	59%
40 weeks	12	2.42 ± 1.38	11	92%

Supplemental Figure S2. Long-term administration of $KBrO_3$ increases tumor number and tumor prevalence in wild-type mice. Number of tumors in the small intestines of 0.15% $KBrO_3$-treated wild-type mice. In the box-whisker plot; blue indicates 16 weeks and red indicates 40 weeks of administration (p = 0.003, Wilcoxon test). The number of mice examined, number of tumors per mouse (mean ± SD), and tumor prevalence (%) are presented in the table.



Supplemental Figure S3. Effect of $KBrO_3$ dose on mutation frequency based on type. Mutation frequency of each mutation type is plotted against different $KBrO_3$ doses. Simple regression analysis was performed for each mutation type and *Mutyh* genotype, with *p*-value adjusted using Bonferroni correction. α = 0.05. ND: not detected. Red line and colored region indicate regression line and 95% confidence interval.



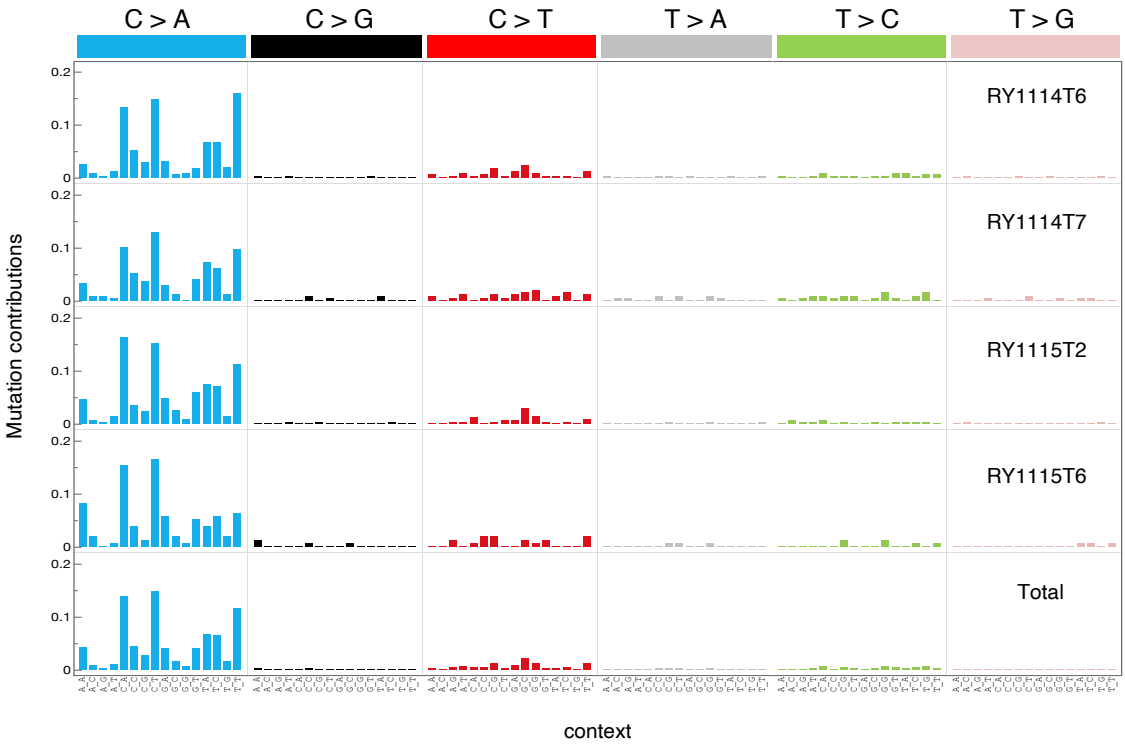
Supplemental Figure S4. Site distribution of base substitution mutations in *rpsL*. Arrows indicate mutated sites and the alphabet inside indicates detected sequences. Promoter region (−35 and −10) and Shine–Dalgarno sequence (SD) in *rpsL* are underlined. The initiation and termination codons are marked with a square.

A) *Mutyh*$^{+/+}$ mice. Upper blue arrow: control, Lower pink arrow: $KBrO_3$-treated.

B) *Mutyh*$^{-/-}$ mice. Upper blue arrow: control, Lower pink arrow: $KBrO_3$-treated.

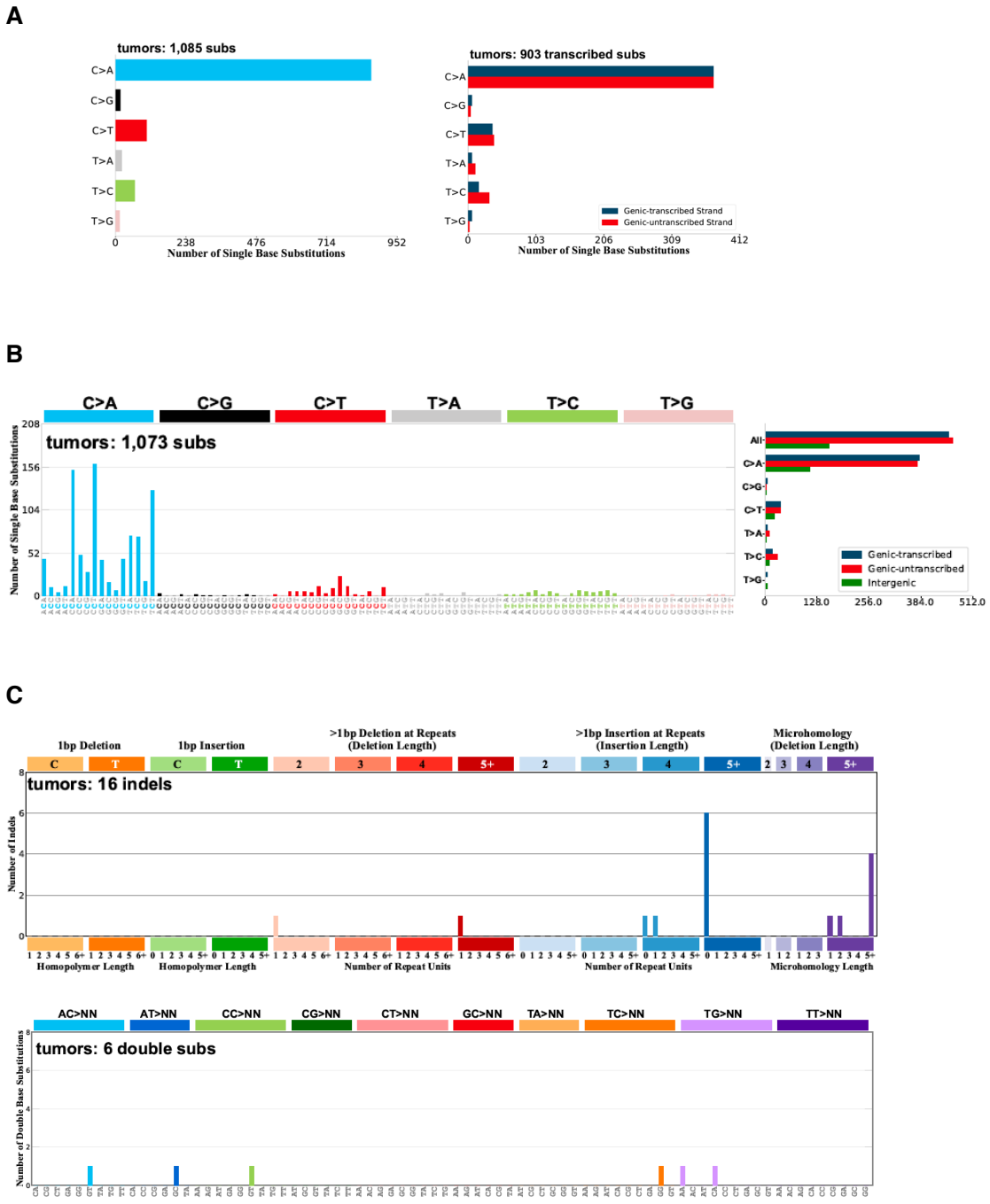
C) Control mice. Upper blue arrow: *Mutyh*$^{+/+}$, Lower pink arrow: *Mutyh*$^{-/-}$.

D) $KBrO_3$-treated mice. Upper blue arrow: *Mutyh*$^{+/+}$, Lower pink arrow: *Mutyh*$^{-/-}$.

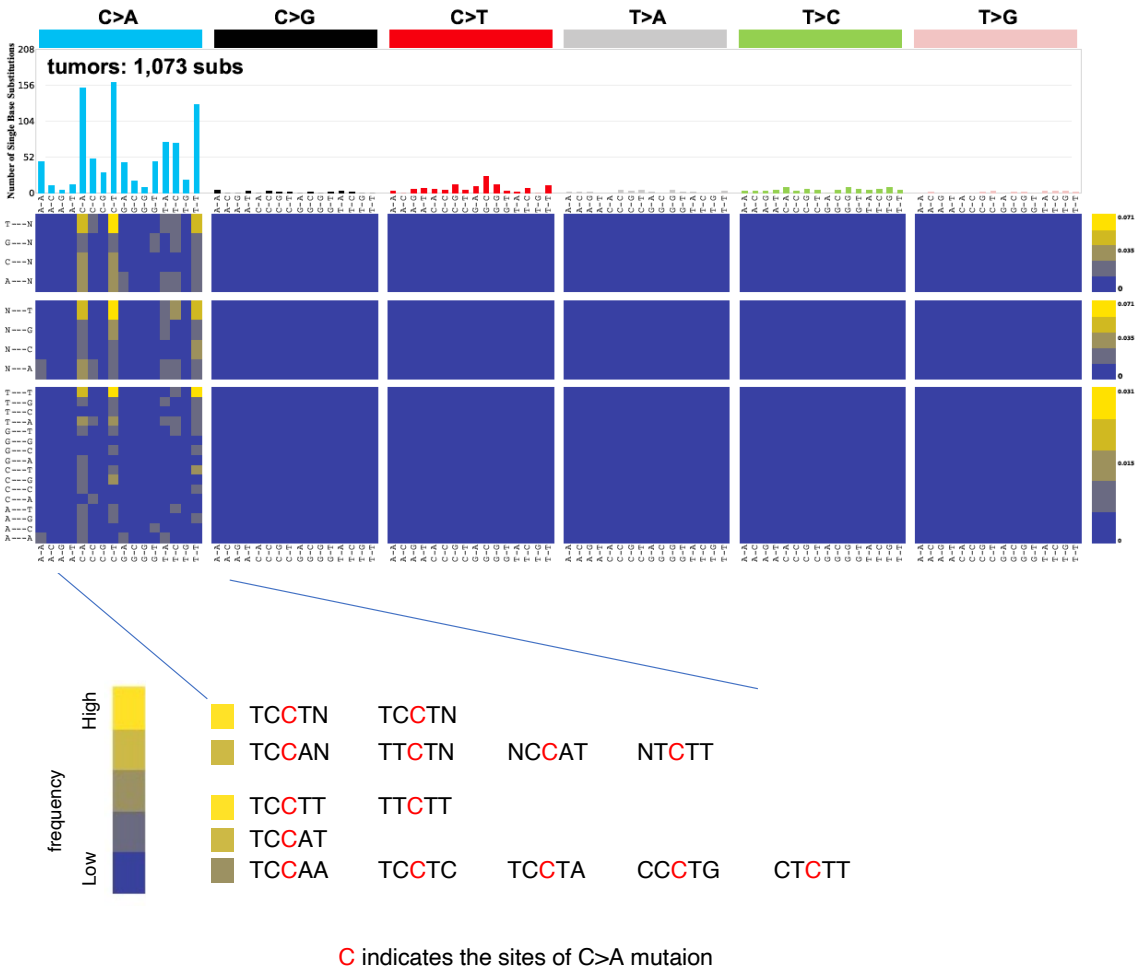


Supplemental Figure S5. Mutation patterns of *Mutyh*$^{-/-}$ tumors.

A) Relative frequency is plotted for each 96-trinucleotide mutation type comprising 5′ and 3″ bases at the mutated site. The tumor ID is written on the right. The sum of all mutations from four tumors is at the bottom (Total).

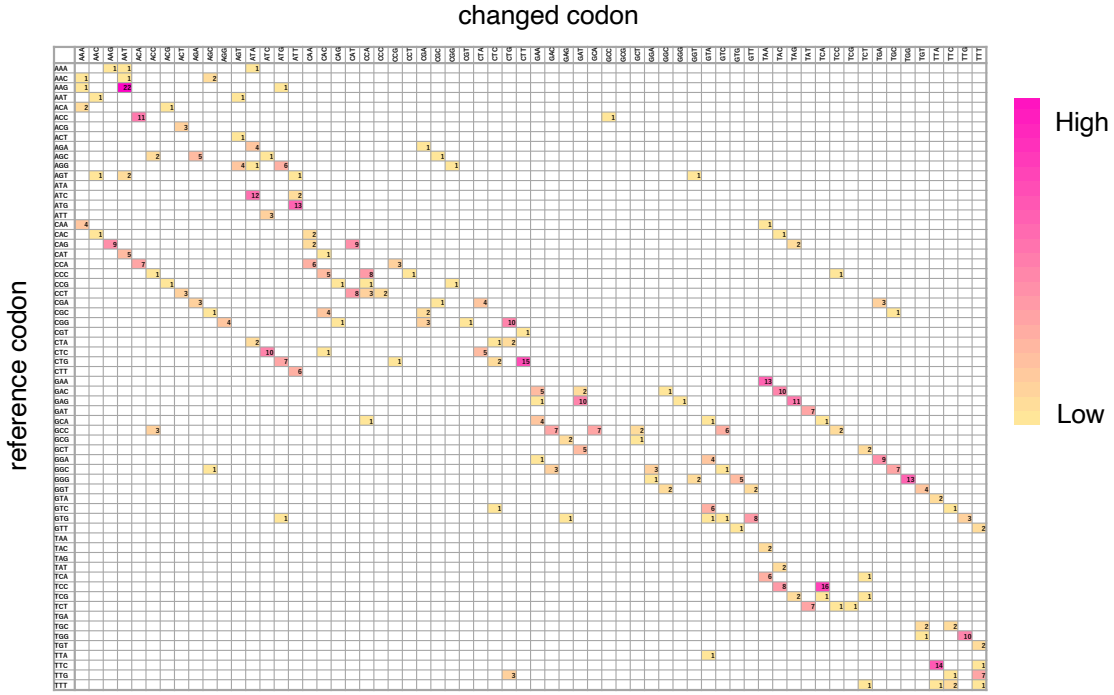


Supplemental Figure S6. Mutational pattern detected using Sigprofiler. The pattern and property of somatic mutations detected in four $KBrO_3$-induced *Mutyh*$^{-/-}$ tumors were analyzed using Sigprofiler. A) Six base substitution type + strand information. B) 96 trinucleotide pattern + strand information. C) Indels and multinucleotide substitution (MNS).



C indicates the sites of C>A mutaion

Supplemental Figure S7. Mutational pattern of *Mutyh*$^{-/-}$ tumor (1536 pentanucleotide). The 1536 pentanucleotide pattern of somatic mutations detected in four $KBrO_3$-induced *Mutyh*$^{-/-}$ tumors was analyzed using Sigprofiler. C>A was frequently detected in the center of the following sequence contexts: TCCTN, TCCTN, TCCTT, and TTCTT.

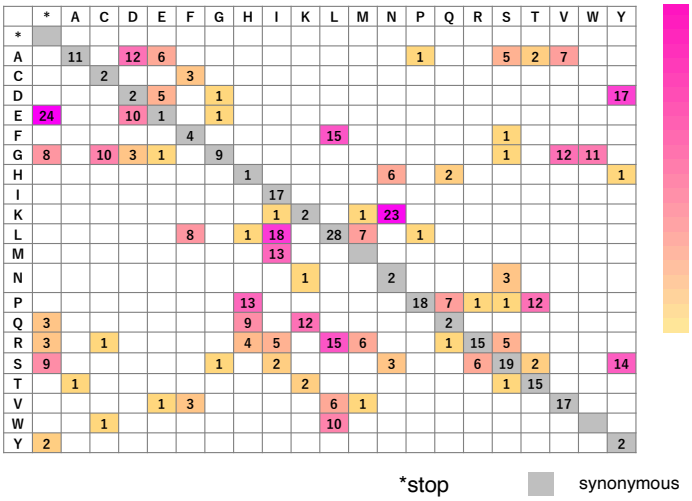


Frequently observed codon changes

codon change	count
AAG > AAT	22
TCC > TCA	16
CTG > CTT	15
TTC > TTA	14
ATG > ATT	13
GAA > TAA	13
GGG > TGG	13
ATC > ATA	12
ACC > ACA	11
GAG > TAG	11
CTC > ATC	10
CGG > CTG	10
GAG > GAT	10
GAC > TAC	10
TGG > TTG	10
CAG > AAG	9
CAG > CAT	9
GGA > TGA	9
CCC > CCA	8
CCT > CAT	8
GTG > GTT	8
TCC > TAC	8

Supplemental Figure S8. Codon changes in the somatic mutations in *Mutyh*$^{-/-}$ tumors. The pattern and frequency of codon changes analyzed using SnpEff are summarized in a table. Rows are reference codons and columns are changed codons (e.g., the number “22” in row “AAG” and column “AAT” indicates how many “AAG” codons have been replaced by “AAT” codons). Frequently observed codon changes are also listed.

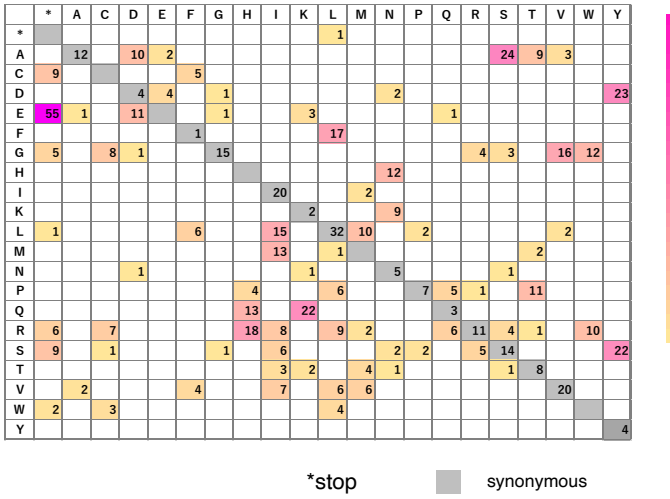
Mutyh$^{-/-}$ mice tumor



Top 14 amino acid changes

amino acid change	count
E > *	24
K > N	23
L > I	18
D > Y	17
F > L	15
R > L	15
S > Y	14
M > I	13
P > H	13
Q > K	12
A > D	12
G > V	12
P > T	12
G > W	11

Human MAP tumor

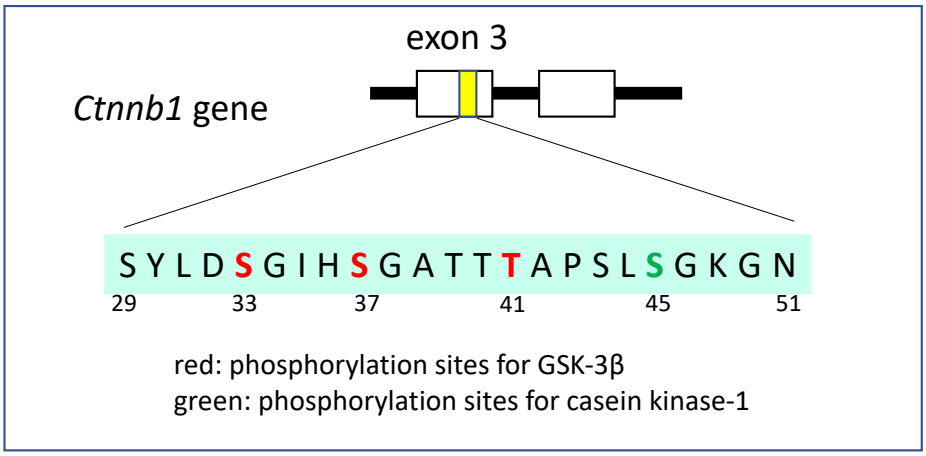


Top 15 amino acid changes

amino acid change	count
E > *	55
A > S	24
D > Y	23
S > Y	22
Q > K	22
R > H	18
F > L	17
G > V	16
L > I	15
M > I	13
Q > H	13
H > N	12
G > W	12
E > D	11
P > T	11

Supplemental Figure S9. Amino acid changes in somatic mutations in *Mutyh*$^{-/-}$ tumors and human MAP tumors. Pattern and frequency of amino acid changes analyzed using SnpEff are summarized in the table. Rows are reference amino acids and columns are changed amino acids (e.g., the number “12” in row “A” and column “D” indicates how many “A” amino acids have been replaced by “D” amino acids). Frequency is represented in the heatmap. The top 10 frequently observed amino acid changes are also listed.

A



B *Ctnnb1* gene mutations in tumors

	mouse ID: RY1114			mouse ID: RY1115		
	normal	tumor	tumor	normal	tumor	tumor
	RY1114 H	RY1114 T6	RY1114 T7	RY1115 H	RY1115 T2	RY1115 T6
33S	chr9:120,950,606	chr9:120,950,606	chr9:120,950,606	chr9:120,950,606	chr9:120,950,606	chr9:120,950,606
	Total count: 44	Total count: 29	Total count: 30	Total count: 10	Total count: 7	Total count: 48
	A : 0	A : 2 (7%, 1+, 1-)	A : 6 (20%, 4+, 2-)	A : 0	A : 0	A : 2 (4%, 1+, 1-)
	C : 44 (100%, 21+, 23-)	C : 27 (93%, 11+, 16-	C : 24 (80%, 11+, 13-)	C : 10 (100%, 8+, 2-)	C : 7 (100%, 5+, 2-)	C : 46 (96%, 27+, 19-
	G : 0	G : 0	G : 0	G : 0	G : 0	G : 0
	T : 0	T : 0	T : 0	T : 0	T : 0	T : 0
	N : 0	N : 0	N : 0	N : 0	N : 0	N : 0
	----------	----------	----------	----------	----------	----------
37S	chr9:120,950,618	chr9:120,950,618	chr9:120,950,618	chr9:120,950,618	chr9:120,950,618	chr9:120,950,618
	Total count: 42	Total count: 26	Total count: 24	Total count: 10	Total count: 6	Total count: 42
	A : 0	A : 1 (4%, 1+, 0-)	A : 0	A : 0	A : 0	A : 2 (5%, 1+, 1-)
	C : 42 (100%, 21+, 21-)	C : 25 (96%, 9+, 16-)	C : 24 (100%, 11+, 13-)	C : 10 (100%, 8+, 2-)	C : 6 (100%, 3+, 3-)	C : 40 (95%, 19+, 21-
	G : 0	G : 0	G : 0	G : 0	G : 0	G : 0
	T : 0	T : 0	T : 0	T : 0	T : 0	T : 0
	N : 0	N : 0	N : 0	N : 0	N : 0	N : 0
	----------	----------	----------	----------	----------	----------
41T	chr9:120,950,630	chr9:120,950,630	chr9:120,950,630	chr9:120,950,630	chr9:120,950,630	chr9:120,950,630
	Total count: 40	Total count: 22	Total count: 28	Total count: 10	Total count: 6	Total count: 42
	A : 0	A : 0	A : 0	A : 0	A : 0	A : 1 (2%, 0+, 1-)
	C : 40 (100%, 21+, 19-)	C : 22 (100%, 9+, 13-)	C : 28 (100%, 15+, 13-)	C : 10 (100%, 7+, 3-)	C : 6 (100%, 3+, 3-)	C : 41 (98%, 21+, 20-
	G : 0	G : 0	G : 0	G : 0	G : 0	G : 0
	T : 0	T : 0	T : 0	T : 0	T : 0	T : 0
	N : 0	N : 0	N : 0	N : 0	N : 0	N : 0
	----------	----------	----------	----------	----------	----------
45S	chr9:120,950,642	chr9:120,950,642	chr9:120,950,642	chr9:120,950,642	chr9:120,950,642	chr9:120,950,642
	Total count: 43	Total count: 25	Total count: 29	Total count: 11	Total count: 10	Total count: 42
	A : 0	A : 0	A : 0	A : 0	A : 0	A : 0
	C : 43 (100%, 21+, 22-)	C : 25 (100%, 12+, 13-)	C : 29 (100%, 16+, 13-)	C : 11 (100%, 8+, 3-)	C : 10 (100%, 7+, 3-)	C : 42 (100%, 21+, 21-)
	G : 0	G : 0	G : 0	G : 0	G : 0	G : 0
	T : 0	T : 0	T : 0	T : 0	T : 0	T : 0
	N : 0	N : 0	N : 0	N : 0	N : 0	N : 0
	----------	----------	----------	----------	----------	----------

Supplemental Figure S10. *Ctnnb1* gene mutations in *Mutyh*$^{-/-}$ tumors. (A) A part of the amino acid sequence (29 to 51) in exon 3 of *Ctnnb1* is summarized. Amino acids in red and green indicate critical sites for phosphorylation by GSK-3 beta or casein kinase-1 (Gao et al.) (30). (B) Whole exome sequencing results of tumor and normal samples at chr9:120,950,606, chr9:120,950,618, chr9:120,950,630, and chr9:120,950,642 corresponding to the codons coding 33S, 37S, 41T, 45S amino acids in *Ctnnb1*, respectively. RY1114H and RY1115H, which are normal tissues, showed that 100% of the sequence reads were the same as the reference sequence, whereas some tumor samples showed that 2%–20% of sequence reads were alternative alleles.

A

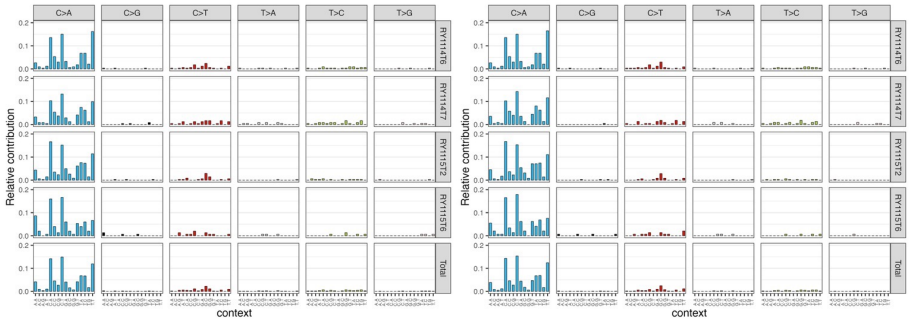
	vcf_mm10_MGPv5	vcf_mm39_MGPv8
recalibration dbSNP	MGPv5_exclude_wild	MGPv8_exclude_wild
recalibration method	Lange et al. (PNAS, 2020)	Lange et al. (PNAS, 2020)
filtering dbSNP	MGPv5_exclude_wild	MGPv8_exclude_wild
vcf for analysis	chr1-19X, no dbSNP, filter PASS, hard filter PASS, not shared	chr1-19X, no dbSNP, filter PASS, hard filter PASS, not shared
number of snv	1073	1060

	RY1114T6	RY1114T7	RY1115T2	RY1115T6	総計
indel	5	9	2	8	24
.	5	9	2	8	24
snv	339	242	342	150	1073
A>C	3	5	2	3	13
A>G	21	23	12	5	61
A>T	6	9	2	3	20
G>A	34	29	26	12	101
G>C	3	4	3	4	14
G>T	272	172	297	123	864
total	344	251	344	158	1097

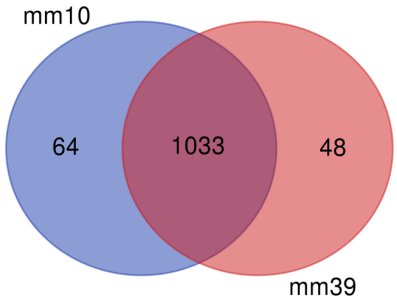
total

	RY1114T6	RY1114T7	RY1115T2	RY1115T6	総計
indel	5	9	2	6	22
.	5	9	2	6	22
snv	339	224	352	145	1060
A>C	3	4	1	1	9
A>G	20	20	10	4	54
A>T	4	5	2	3	14
G>A	36	23	24	12	95
G>C	3	1	3	4	11
G>T	273	171	312	121	877
total	344	233	354	151	1082

total



B



Supplemental Figure S11. Extraction of somatic mutations from tumor WES using mm10 or mm39 reference genome. A) The number of mutations detected from each tumor and the 96-trinucleotide mutation pattern shows similarity between mm10 and mm39. C) The number of variants common or specific in the data from mm10 and mm39.