

## **Supplemental Materials and Methods**

### **Statistical analysis of an independent cohort**

The data set from an independent cohort (1) was utilized for the evaluation of *HOTAIR* expression as a prognostic factor in CRC. The microarray gene expression data and clinical information of the 320 CRC tissue samples were available at the GEO database with accession code GSE24551. RMA normalization (2) for each probe set was performed with Affymetrix Power Tools software, and the default options of RMA (with background correction, quantile normalization, and log-transformation) were used. On the probe sets of the Affymetrix Human Exon 1.0 ST Array, 10 probes were mapped to the *HOTAIR* region, and the probe (probeset\_ID: 3456485, position: chr12: 54,360,111-54,360,139) with the largest variance among the 320 samples was selected as one corresponding to *HOTAIR*. We then defined the *HOTAIR* expression level as “*HOTAIR* high” if it was higher than the median expression value of *HOTAIR*, and the term “*HOTAIR* low” was defined for *HOTAIR* expression levels lower than the median.

### **Supplementary References**

1. Sveen A, Agesen TH, Nesbakken A, Rognum TO, Lothe RA, Skotheim RI. Transcriptome instability in colorectal cancer identified by exon microarray analyses: Associations with splicing factor expression levels and patient survival. *Genome Med* 2011; 3: 32.
2. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 2003; 31: e15.

### **Supplementary Figure Legends**

#### Figure S1

Kaplan-Meier overall survival curves according to *HOTAIR* level in an independent CRC subgroup (n = 320). The overall survival rate of the high *HOTAIR* expression group (n = 160) was significantly higher than that of the low expression group (n = 160; logrank test;  $P < 0.0001$ ). (See Supplemental Material)

#### Figure S2

Heatmap representing expression values for 32 CRC samples enriched in signatures of *HOTAIR*-induced SUZ12, H3K27me3, EZH2, and PRC2 occupancy, respectively. The red and blue colors indicate high and low expression, respectively.

Figure S3

A. *HOTAIR/GAPDH* expression levels determined by quantitative real time PCR analysis of the eight CRC cell lines and MCF7 (breast cancer cell line). Values are expressed relative to *HOTAIR* abundance in MCF7, a cell line that expresses endogenous *HOTAIR*.

B. *HOTAIR/GAPDH* expression levels by quantitative real time PCR analysis of HCT116 after infection with *HOTAIR* expression lentiviral vector or empty vector. Error bars represent S.D. (n=3)

C. *HOTAIR/GAPDH* expression levels by quantitative real time PCR analysis of SW480 after transfection with siRNA targeting *HOTAIR* (two individual siRNAs; si HOTAIR1, si HOTAIR2) and negative control siRNA (si n.c.). Error bars represent S.D. (n=3)