

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- ☐ ☒ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- ☐ ☒ A description of all covariates tested
- ☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☐ ☒ For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☐ ☒ Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

We enrolled a total of 647 subjects who underwent surgery or endoscopy of the lower gastrointestinal tract at Kyoto University Hospital, Hyogo College of Medicine Hospital, and Cancer Institute Hospital. Among these, 76 had been diagnosed as ulcerative colitis and the remaining 571 were diagnosed as other disease such as colorectal polyp or cancer without colitis, or were turned out to be healthy individuals. This study was approved by the ethical committees at Kyoto University, Hyogo College of Medicine, and Cancer Institute Hospital. Informed consent was obtained from all participants. As for external data sets, whole exome sequencing (WES) data of paired tumor/germline control samples from CRC ($n=356$) patients were downloaded from the TCGA Data Portal (<https://portal.gdc.cancer.gov/>). Four previously published mutation lists of WES obtained from 65 colitis-associated cancer in UC individuals were also analysed.

Data analysis

Detailed descriptions of the software and analysis have been provided in Online Methods. Sequencing data was processed using our in-house pipeline Genomon2. External bam files were converted to fastq format using biobambam2. Copy number analysis using sequencing data was performed by in-house pipeline CNACS. Significantly mutated genes were identified using dNdScv and MutSigCV. Mutational signature was evaluated using pmsignature and deconstructSigs. Clonal architecture was estimated by PyClone (0.13.0). Phylogenetic analysis was performed by using MEGA X (v10.0.5). Transcriptome analysis was performed by STAR (v2.5.3.), featureCounts (v1.5.3.), edgeR (v3.24.3), and GSEA (v2.2.4). Statistical analyses were performed using R (3.5.1). Images were analyzed by using ImageJ (1.52a).

List of programs and softwares:

Genomon2 pipeline: version 2.6.0 (<https://genomon.readthedocs.io/ja/latest>)
 - Burrows-Wheeler Aligner: version 0.7.10 (<https://sourceforge.net/projects/bio-bwa/>)
 - picard-tools: version 1.39 (<http://picard.sourceforge.net/>)
 - biobambam2: version 2.0.85 (<https://www.sanger.ac.uk/science/tools/biobambam>)
 - GenomonMutationFilter: (<https://github.com/Genomon-Project/GenomonMutationFilter>)
 - EBCall: version 2 (<https://github.com/friend1ws/EBCall>)
 Integrative Genomics Viewer (IGV): version 2.4.6 (<http://software.broadinstitute.org/software/igv/>)

MutSigCV: version 1.4. (<http://software.broadinstitute.org/cancer/software/genepattern/modules/docs/MutSigCV>)
 dNdScv: version 0.0.1 (<https://github.com/im3sanger/dndscv>)
 CNACS: (http://plaza.umin.ac.jp/kyoto_tumorpatho/CNACS.html)
 pmsignature: version 0.3.0 (<https://github.com/friend1ws/pmsignature>)
 deconstructSigs: version 1.8.0 (<https://github.com/raerose01/deconstructSigs>)
 PyClone: version 0.13.0 (<https://github.com/arothe85/pyclone>)
 MEGA X: version 10.0.5 (<https://www.megasoftware.net/home>)
 STAR: v2.5.3 (<https://github.com/alexdobin/STAR>)
 featureCounts: v1.5.3 (<http://bioinf.wehi.edu.au/featureCounts/>)
 edgeR: 3.24.3 (<https://bioconductor.org/packages/release/bioc/html/edgeR.html>)
 GSEA: v2.2.4 (<http://software.broadinstitute.org/gsea/index.jsp>)
 R: v.3.5.1 (<https://cran.r-project.org/>)
 ImageJ: v1.52a (<https://imagej.nih.gov/ij/index.html>)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All the WES and RNA-seq data have been deposited in the European Genome-Phenome Archive (<http://www.ebi.ac.uk/ega/>) under accession numbers EGAS00001003801 and EGAS00001003802, respectively.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to determine sample size since this is an exploratory study. We enrolled as many patients as possible who provided consent for our study during the enrollment period between 1/Oct/2014 and 31/Mar/2019. Cryo-preserved colorectal cancer samples were provided from Cancer Institute Hospital. In total, we collected 1568 samples from 76 UC patients and 571 healthy or non-UC individuals. The sample size of our study was sufficient to confirm novel findings with statistical significance. As for external data sets, 356 sporadic CRCs (sCRCs) and 65 colitis-associated cancers were also analysed.
Data exclusions	For the analysis of TCGA sCRC samples, we excluded whole-genome amplified samples to avoid artefactual mutation callings. We also excluded hypermutated TCGA sCRCs from the analysis of comparison of driver genes between UC non-dysplasia and TCGA sCRCs because high background mutation rate of the tumours might make the difference difficult to detect. We also enrolled only non-hypermutated sCRCs from the Japanese cohort for the same reason.
Replication	Three or more biological replicates were generated for every assay, except for immunoblot experiments where two biological replicates were generated and the same results were confirmed among two replicates. We tested significant difference for every assay as described in the figure legends and Methods.
Randomization	Not applicable since this is a case-series study which was therefore not planned to detect any difference in effects between the cohorts with and without intervention.
Blinding	Two expert pathologists, blinded to clinical information, reviewed all the specimens obtained from UC patients. Researchers were blinded to genotype information at measurement of each animal experiment.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☐	☒ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Monoclonal anti-c-Myc (supplier name, Sigma-Aldrich; catalogue number, M4439; clone name, 9E10; lot number, 026M4825V; dilution, 1/2000) Anti-mouse IgG HRP (supplier name, GE Healthcare; catalogue number, NA9310-1ML; lot number, 16919626; dilution, 1/5000) Polyclonal anti-ZC3H12A (supplier name, Atlas Antibodies; catalogue number, HPA032053; lot number, R32479; dilution, 1/500) Polyclonal anti-Flag (supplier name, Sigma-Aldrich; catalogue number, F7425; lot number, 078M4886V; dilution, 1/2000) Monoclonal anti-beta-Actin (supplier name, Santa Cruz; catalogue number, sc-47778 HRP; clone name, C4; lot number, K1418; dilution, 1/1000) Anti-rabbit IgG HRP (supplier name, GE healthcare; catalogue number, NA9340V; lot number, 16926347; dilution, 1/5000)
Validation	Monoclonal anti-c-Myc (supplier name, Sigma-Aldrich; catalogue number, M4439) – immunoblotting – human cells; Validation: Manufacturer - https://www.sigmaaldrich.com/catalog/product/sigma/m4439 Polyclonal anti-ZC3H12A (supplier name, Atlas Antibodies; catalogue number, HPA032053) – immunoblotting – human cells; Validation: Manufacturer - https://www.atlasantibodies.com/products/antibodies/primary-antibodies/triple-a-polyclonals/zc3h12a-antibody-hpa032053/ Polyclonal anti-Flag (supplier name, MERCK; catalogue number, F7425) – immunoblotting – human cells; Validation: Manufacturer - https://www.sigmaaldrich.com/catalog/product/sigma/f7425 Monoclonal anti-beta-Actin (supplier name, Santa Cruz; catalogue number, sc-47778 HRP) – immunoblotting – human cells; Validation: Manufacturer - https://datasheets.scbt.com/sc-47778.pdf

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	HT29, SW480, DLD1, HCT116, LoVo, and HeLa cells were obtained from ATCC.
Authentication	None of the cell lines were authenticated.
Mycoplasma contamination	All cell lines used for this study were tested negative for mycoplasma using Myco Alert Mycoplasma Detection Kit (Lonza).
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used in this study.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	VilCre mice (004586) and R26-LacZ mice (003309) were obtained from The Jackson Laboratory. Nfkbiz flox mice (RBRC06410) were obtained from Riken BRC. VilCreER mice were kindly gifted by Sylvie Robine. All mice used in this study were C57BL/6 strains. In the analysis of clonal expansion in chronic DSS colitis, we analysed 11 R26-LacZ/wt;Nfkbiz wt/wt;VilCreER (control) mice (5 male and 6 female) and 10 R26-LacZ/wt;Nfkbiz fl/fl;VilCreER (Nfkbiz cKO) mice (5 male and 5 female) for the pre-colitis assessment. We also analysed 20 control mice (13 male and 7 female) and 21 Nfkbiz cKO mice (14 male and 7 female) for the post-colitis assessment. All mice were injected with tamoxifen at 6 weeks of age. In the transcriptome analysis of DSS-induced colitis, 3 male Nfkbiz fl/fl (control) mice and 3 male Nfkbiz fl/fl;VilCre (Nfkbiz cKO) mice were analysed at 8 weeks of age. In the AOM-DSS tumour model, we started experiments with 25 Nfkbiz fl/fl (control) mice (18 male and 7 female) and 25 Nfkbiz fl/fl;VilCre (Nfkbiz cKO) mice (18 male and 7 female) at 8 weeks of age.
Wild animals	This study did not involve wild animals
Field-collected samples	This study did not involve field-collected samples
Ethics oversight	All in vivo experiments in animals were approved by Kyoto University Animal Care and Use Committee.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	Participants were recruited through at Kyoto University, Hyogo College of Medicine, and Cancer Institute Hospital. We enrolled a total of 647 patients who underwent surgery or endoscopy of the lower gastrointestinal tract at Kyoto University Hospital, Hyogo College of Medicine Hospital, and Cancer Institute Hospital. Among these, 76 had been diagnosed as ulcerative colitis (UC) and the remaining 571 (healthy or non-UC) were diagnosed as other disease such as colorectal polyp or cancer without associated colitis or turned out to be healthy individuals. The detailed characteristics of cases were provided in Supplementary Table 1.
Recruitment	Patients who underwent surgery or endoscopy of the lower gastrointestinal tract between 1/Oct/2014 and 31/Mar/2019 at Kyoto University and Hyogo College of Medicine were recruited without any selection. Cryo-preserved non-hypermutated sCRC samples between 1/Dec/2006 and 31/May/2016 were provided from Cancer Institute Hospital without any selection.
Ethics oversight	This study was approved by the ethical committees at Kyoto University, Hyogo College of Medicine, and Cancer Institute Hospital. Informed consent was obtained from all participants.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration	Not applicable. This study is not a clinical trial.
Study protocol	Not applicable. This study is not a clinical trial.
Data collection	Not applicable. This study is not a clinical trial.
Outcomes	Not applicable. This study is not a clinical trial.