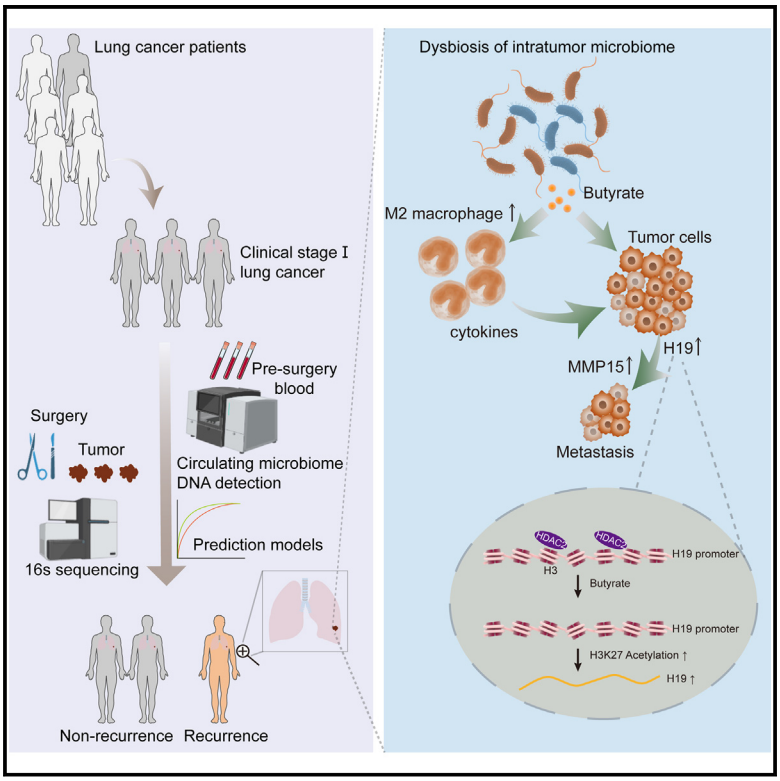
Cell Reports
Medicine

Article

Intratumor microbiome-derived butyrate promotes lung cancer metastasis

Graphical abstract



Highlights

- Intratumor microbiomes are associated with recurrence of lung cancer
- Intratumor microbiome and circulating microbiome DNA can predict cancer recurrence
- Intratumor-microbiome-derived butyrate promotes lung cancer metastasis

Authors

Yi Ma, Haiming Chen, Haoran Li, ..., Yun Li, Jie Wang, Mantang Qiu

Correspondence

surgeonli@hotmail.com (Y.L.), wangjie5204007@hotmail.com (J.W.), qiumantang@163.com (M.Q.)

In brief

Ma et al. illustrate the potential of intratumor microbiome signatures and circulating microbiome DNA as promising biomarkers for lung cancer recurrence. Butyrate-producing bacteria, such as *Roseburia*, are enriched in lung cancer with early recurrence, and intratumormicrobiome-derived butyrate upregulates H19 expression and facilitates M2 macrophage polarization, thereby promoting metastasis.

Ma et al., 2024, Cell Reports Medicine *5*, 101488
April 16, 2024
https://doi.org/10.1016/j.xcrm.2024.101488

CellPress

Cell Reports Medicine

CellPress
OPEN ACCESS

Article

Intratumor microbiome-derived butyrate promotes lung cancer metastasis

Yi Ma,[1,2,9] Haiming Chen,[1,3,9] Haoran Li,[1,3,9] Meiling Zheng,[4,5,9] Xianglin Zuo,[6] Wenxiang Wang,[1,3] Shaodong Wang,[1,3] Yiming Lu,[7] Jun Wang,[8] Yun Li,[1,3,*] Jie Wang,[6,*] and Mantang Qiu[1,3,10,*]

[1]Department of Thoracic Surgery, Thoracic Oncology Institute, Peking University People's Hospital, Beijing 100044, China
[2]Department of Thoracic Surgery, Shanghai Pulmonary Hospital, Shanghai 200433, China
[3]Institute of Advanced Clinical Medicine, Peking University, Beijing 100191, China
[4]Dongzhimen Hospital, Beijing University of Chinese Medicine, Beijing 100010, China
[5]Department of Chinese Medicine, Peking University People's Hospital, Beijing 100044, China
[6]Department of Science & Technology, Biobank of Jiangsu Cancer Hospital, Jiangsu Cancer Hospital, Jiangsu Institute of Cancer Research, The Affiliated Cancer Hospital of Nanjing Medical University, Nanjing 210009, China
[7]Department of Genetics & Integrative Omics, State Key Laboratory of Proteomics, National Center for Protein Sciences, Beijing Institute of Radiation Medicine, Beijing 100850, China
[8]CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China
[9]These authors contributed equally
[10]Lead contact
*Correspondence: surgeonli@hotmail.com (Y.L.), wangjie5204007@hotmail.com (J.W.), qiumantang@163.com (M.Q.)
https://doi.org/10.1016/j.xcrm.2024.101488

SUMMARY

Most recurrences of lung cancer (LC) occur within 3 years after surgery, but the underlying mechanism remains unclear. Here, we collect LC tissues with shorter (<3 years, recurrence group) and longer (>3 years, non-recurrence group) recurrence-free survival. By using 16S sequencing, we find that intratumor microbiome diversity is lower in the recurrence group and butyrate-producing bacteria are enriched in the recurrence group. The intratumor microbiome signature and circulating microbiome DNA can accurately predict LC recurrence. We prove that intratumor injection of butyrate-producing bacteria *Roseburia* can promote subcutaneous tumor growth. Mechanistically, bacteria-derived butyrate promotes LC metastasis by increasing expression of H19 in tumor cells through inhibiting HDAC2 and increasing H3K27 acetylation at the H19 promoter and inducing M2 macrophage polarization. Depletion of macrophages partially abolishes the metastasis-promoting effect of butyrate. Our results provide evidence for the cross-talk between the intratumor microbiome and LC metastasis and suggest the potential prognostic and therapeutic value of the intratumor microbiome.

INTRODUCTION

Lung cancer accounts for the highest cancer-related mortality around the world.[1,2] Despite encouraging results achieved with targeted therapy, immunotherapy, and surgery in recent years, the 5-year overall survival is less than 20%, mainly as a result of late diagnosis and the high frequency of recurrence and metastasis.[3] Patients with lung cancer usually suffer recurrence within 3 years after surgery.[4] Therefore, exploring the mechanisms related to recurrence and metastasis and identifying diagnostic biomarkers and potential therapeutic targets are important tasks related to lung cancer.

An aberrant microbiome has been one of new hallmarks of cancer and is deeply involved in the process of cancer progression, while the biological role of a specific bacteria might be context specific.[5,6] Bacteria belonging to the genus *Lactobacillus* are well-known probiotics, but recent evidence suggests that indole-producing bacteria, including *Lactobacillus murinus*, promote pancreatic carcinoma growth by activating the aryl hydrocarbon receptor pathway.[7] For lung cancer, it has been reported that the commensal microbiome contributes to tumor cell proliferation by cross-talk with myeloid cells and γδT cells. In our previous study, we also found that lung cancers with different radiological features have distinct microbiome composition and that alpha diversity is higher in the lung cancer subtype with indolent clinical behavior.[8] Identification of novel bacterium related to lung cancer prognosis and the underlying molecular mechanism may lead to new diagnostic and therapeutic strategies to improve survival outcomes for lung cancer.

In the current study, we investigated microbiome composition of early-stage lung cancer with short or long recurrence-free survival (RFS) by 16S rRNA sequencing. Intratumor microbiome diversity was reduced in patients with short RFS, and butyrate-producing bacteria were enriched in patients with lung cancer with short RFS. The intratumor microbiome signature and pre-operative circulating microbiome DNA (cmDNA) could accurately

Table 1. Clinical characteristics of patients in 16S cohort

Clinical characteristics	R (n = 29)	NR (n = 29)	p value
Recurrence-free survival (years), mean ± SD	1.31 ± 0.83	5.01 ± 0.71	<0.001
Age (years), mean ± SD	59.14 ± 8.23	61.41 ± 9.65	0.338
Gender			0.793
Female	15	14	–
Male	14	15	–
BMI (kg/m^2), mean ± SD	24.41 ± 2.57	24.37 ± 3.47	0.958
Smoking history			0.188
Yes	13	18	–
No	16	11	–
Tumor diameters (cm), mean ± SD	2.19 ± 0.54	1.94 ± 0.67	0.127
Clinical T stage			0.242
T1a	1	4	–
T1b	12	14	–
T1c	16	11	–
TNM stage			<0.001
I	10	29	–
II	5	0	–
III	14	0	–
Pathology			0.611
LUAD	26	28	–
LUSC	2	1	–
Others	1	0	–

R, recurrence group; NR, non-recurrence group; BMI, body mass index; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; Others, lung adenosquamous carcinoma.

predict lung cancer recurrence. We found that *Roseburia*, a well-known butyrate-producing bacteria, could promote lung cancer metastasis, and butyrate could directly promote lung cancer cell invasion by increasing H19 and MMP15 expression, and induce M2 macrophage polarization.

RESULTS

Tumor microbial diversity is associated with recurrence in patients with resected lung cancer

To explore the association between intratumor lung tumor microbiome composition and recurrence of lung cancer, we first established an early-stage lung cancer cohort including patients who suffered recurrence within 3 years after surgery (recurrence [R] group, median RFS 1.3 years) and clinical stage-matched long-term survivors who survived more than 3 years without recurrence (non-recurrence [NR] group, median RFS 4.9 years) from our center (Figure S1A, 16S cohort). Patients in the R or NR group were matched with respect to age, gender, BMI, smoking history, clinical stage, tumor diameters, and pathology (Table 1). In addition, the TNM stage was more advanced in patients in the R group than those in the NR group, which is consistent with the characteristics of tumor recurrence. Bacterial DNA was extracted from 58 patients with paired surgically resected lung cancer tumor and adjacent normal lung tissue (29 R and 29 NR), and taxonomic profiling via 16S rRNA gene sequencing was performed. The species accumulation curve indicated that the sequencing depth is adequate (Figure S1B).

We first measured the tumor microbial diversity using different methodologies (observed taxonomic units, Shannon and Simpson indices) and found that the alpha diversity of the tumor microbiome, defined as the abundance and diversity of species present within each tumor sample, was significantly higher in NR patients compared to R patients ($p = 0.002$ for Shannon and $p = 0.001$ for Simpson, Figure 1A). Similarly, the alpha diversity of the normal lung microbiome was higher in NR patients compared to R patients ($p = 0.001$ for Shannon and $p = 0.004$ for Simpson, Figure S1C). We then explored the relationship between lung cancer tumor microbial diversity and RFS by stratifying the patients in the two groups based on median diversity obtained by Shannon index. As expected, we found that patients with low alpha diversity have significantly decreased RFS compared to those with high alpha diversity (hazard ratio [HR] = 2.658, 95% confidence interval [CI]: 1.233–5.732, $p = 0.0095$, Figure 1B). Our findings indicate that the tumor alpha diversity could serve as a predictor for RFS in patients with resected lung cancer, suggesting that the intratumor microbiome may be involved in lung cancer recurrence.

To further extend our understanding of the role of microbiome diversity and its association with recurrence, we compared the overall microbiome composition in tumors between the R and NR groups using microbial beta diversity. Principal-coordinate analysis (PCoA) using Bray-Curtis distance showed clearly different clusters between the R and NR groups (Figure 1C). Analysis of similarity (ANOSIM) was performed to evaluate the statistical difference between the microbiome compositions. We found a significant difference of overall tumor microbiome composition between the R and NR groups ($R = 0.205$, $p = 0.001$, Figure 1C). In addition, in normal lung tissues, a similar difference of overall microbiome composition was also found between the R and NR groups (ANOSIM $R = 0.108$, $p = 0.003$, Figure S1D).

Tumor microbiome communities are significantly different between R and NR patients

Considering the relationship between diversity and RFS in patients with lung cancer, we next sought to identify the differences in microbial communities between R and NR patients. At the phylum level, *Firmicutes*, *Bacteroidetes*, *Proteobacteria*, and *Actinobacteria* were mainly compositions in both the R and NR groups, regardless of whether they were tumor or normal lung tissues (Figures 2A and S1E). Furthermore, the proportions of *Firmicutes* were increased, but the proportions of *Proteobacteria* and *Actinobacteria* were decreased, in tumor and normal tissues of R patients compared those that in NR patients (Figures 2B and S1F). At the genus level, *Faecalibacterium*, *Roseburia*, *Blautia*, and *Bifidobacterium* showed relatively high abundance in both groups (Figures 2A and S1G). Compared to the NR group, the proportions of *Roseburia* and *Blautia* were increased, while the proportions of *Faecalibacterium* and *Bifidobacterium* were decreased, in the R group (Figures 2A and S1H).

In order to further identify the differential microbiota between the R and NR groups and find potential microbial biomarkers,

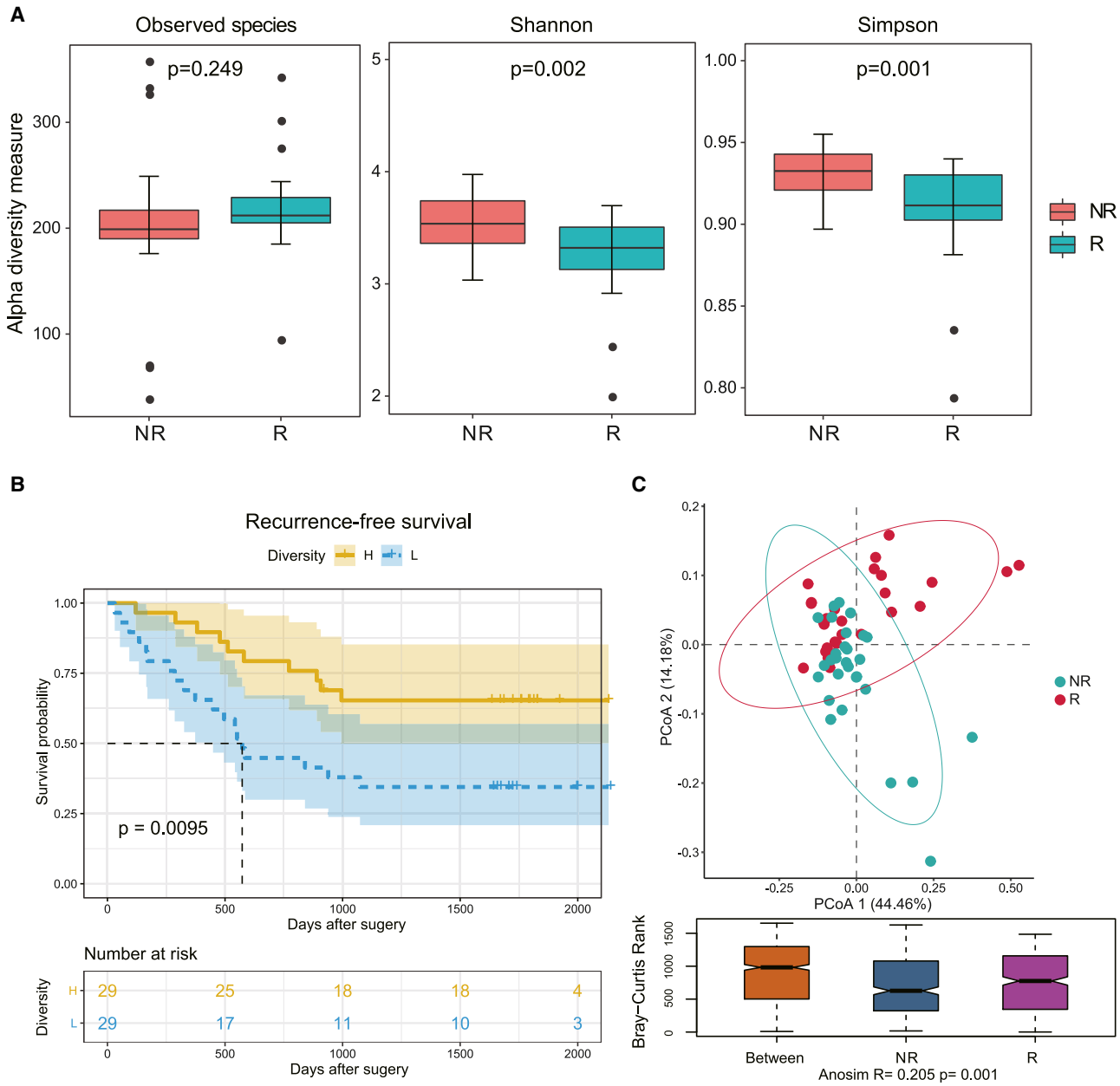


Figure 1. Intratumor microbial diversity correlates with recurrence of patients with lung cancer

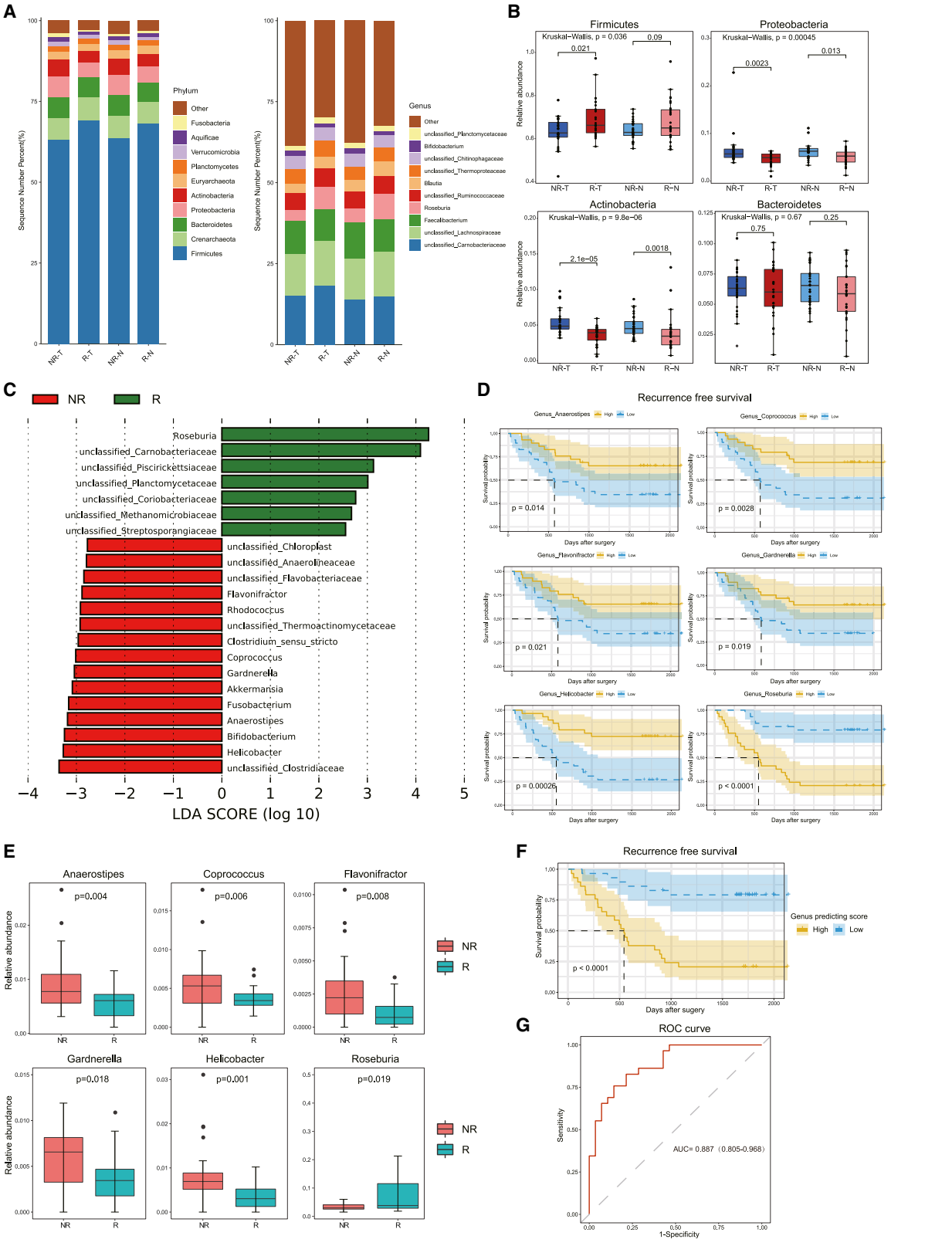
(A) Alpha diversity in R and NR groups (observed species, Shannon, and Simpson indices).

(B) Kaplan-Meier plot of patients with lung cancer defined by alpha diversity.

(C) PCoA using Bray-Curtis metric distances of beta diversity.

R, recurrence group; NR, non-recurrence group; H, high diversity; L, low diversity; PCoA, principal coordinate analysis. The error bars indicate the standard deviations. See also Figure S1.

we performed linear discriminant analysis of effect size (LEfSe) at the genus level. In tumor tissues, LEfSe revealed 22 features that may discriminate between the R and NR groups (Figure 2C). *Roseburia* was the most evident biomarker for the R group. We also retrieved another cohort according to the same criteria at Jiangsu Cancer Hospital (baseline clinical characteristics are shown in Table S1). RT-PCR assays confirmed that *Roseburia* was more abundant in the R group (Figure S1I). *Bifidobacterium*, *Helicobacter*, and *Akkermansia* were biomarkers for the NR group. Using the Lasso regression model for further selection, we found six genera (*Roseburia*, *Helicobacter*, *Gardnerella*, *Flavonifractor*, *Coprococcus*, and *Anaerostipes*) that may serve as potential biomarkers to discriminate between the R and NR groups. These six genera were significantly correlated with RFS (Figure 2D). The relative abundances of these six genera were significantly different between the R and NR groups



(legend on next page)

(Figure 2E), and genus predicting scores were then calculated based on the proposed formula (Table S2). Patients were then divided into high- and low-risk groups according to the median predicting score. The Kaplan-Meier survival curve revealed that RFS in the high-risk group was significantly shorter (HR = 6.543, 95% CI: 2.640–16.217, $p < 0.001$, Figure 2F). Importantly, the genus predicting score was still an independent predictor of RFS in the multivariate Cox regression model (HR = 5.549, 95% CI: 2.669–11.535, $p < 0.001$, Table S3) after adjusting for the TNM stage. The area under the curve (AUC) value of the genus predicting score was 0.887 (Figure 2G).

LEfSe was also performed at the genus level to identify the differential microbiota in normal lung tissue between the R and NR groups. In normal lung tissues, LEfSe revealed 22 features that may discriminate between the R and NR groups (Figure S2A). Consistent with the findings in tumor tissue, *Roseburia* was the most evident biomarker for the R group, and *Bifidobacterium*, *Helicobacter*, and *Akkermansia* were biomarkers for the NR group. In addition, the butyrate-producing bacteria *Butyricicoccus* and *Butyricimonas* were also features for the R group. Seven genera (*Akkermansia*, *Helicobacter*, *Butyricicoccus*, *Gp6*, *Leptotrichia*, *Rhodococcus*, and *Roseburia*) were further identified to discriminate between the R and NR groups using the Lasso regression model. Six of these genera were significantly correlated with RFS (Figure S2B). Differential abundances of these seven genera in the R and NR groups were found (Figure S2C). The genus predicting scores were then calculated based on the proposed formula (Table S2). Patients were then divided into high- and low-risk groups according to the median predicting score. RFS was significantly shorter in the high-risk group (HR = 3.731, 95% CI: 1.642–8.475, $p = 0.002$, Figure S2D). Importantly, the genus predicting score was an independent predictor of RFS in the multivariate Cox regression model (HR = 2.456, 95% CI: 1.280–4.710, $p = 0.007$, Table S4) after adjusting for the TNM stage. The AUC value of the genus predicting score was 0.859 (Figure S2E).

cmDNA signatures are promising biomarkers for lung cancer recurrence

Recently, cmDNA emerged as promising biomarkers in cancer diagnosis.[9] Therefore, we recruited another cohort to explore the relation between cmDNA and lung cancer recurrence (Figure 3A, cmDNA cohort). A total of 63 patients with lung cancer were enrolled, and their plasma was sequenced through whole-genome sequencing. Patients in the R or NR group showed no differences in age, gender, BMI, smoking history, or pathology (Table S5). At the genus level, *Acinetobacter*, *Cutibacterium*, *Comamonas*, and *Staphylococcus* showed relatively higher abundance in both groups (Figure S3A). As expected, we found 22 shared genera in both plasma and tumor tissues of R and NR patients, including *Helicobacter* and *Gardnerella*, which were biomarkers in tumor tissues (Figure 3B; Table S6). Consistent to some extent with tumor tissues, alpha diversity of the cmDNA was higher in NR patients compared to R patients, although no statistical difference was found ($p = 0.74$ for Shannon and $p = 0.33$ for Simpson, Figure 3C). Moreover, PCoA using Bray-Curtis distance showed a difference in cmDNA between the R and NR groups (ANOSIM $R = 0.078$, $p = 0.018$, Figure 3D).

The patients were then randomly split into a discovery set (70%) and a validation set (30%) for model calibration and validation (Figure 3A). Seven genera with predictive probability were found using MaAslin2 (Table S7). *Klebsiella*, *Massilia*, and *Microbulbifer* were enriched in the R group, while *Cutibacterium*, *Comamonas*, *Staphylococcus*, and *Hydrogenophilus* were enriched in the NR group (Figures S3B and S3C). Principal-component analysis showed clearly different clusters between the R and NR groups using these seven genera (Figures 3E and S3D). A random forest model based on these seven genera was constructed and reached an AUC of 0.857 in the discovery set (Figure 3F). We used the validation set to further evaluate the model and reached acceptable predictive power (AUC = 0.717, Figure 3F), albeit lower than that in the discovery set. In addition, we also found that genus including *Staphylococcus*, *Massilia*, and *Klebsiella* were significantly correlated with RFS (Figure 3G). These results suggested that cmDNA signatures were promising non-invasive biomarkers for predicting recurrence in early-stage lung cancer before surgery.

Microbiota-derived butyrate promotes progression of lung cancer at low concentrations

Given that butyrate-producing bacteria, especially *Roseburia*, were enriched in both tumor and normal tissues in the R group, we further explored whether *Roseburia* can enhance lung cancer progression. Using BLAST, we found that the sequence of *Roseburia* was most closely aligned with *Roseburia intestinalis*, which was chosen for further experiments. Mice with subcutaneous xenograft lung cancer were intratumorally injected with *R. intestinalis* (*Roseburia* group) or PBS every 3 days. Introduction of *R. intestinalis* into the subcutaneous xenografts in mice increased tumor volume ($p < 0.01$; Figure 4A). Furthermore, lung metastasis was only found in the *Roseburia* group (Figure S4A).

Various butyrate-producing bacteria such as *Roseburia* and *Butyricicoccus* were disturbed in the R group, indicating that butyrate may play a role in tumor recurrence and metastasis. Previous studies have demonstrated that butyrate may play

Figure 2. Intratumor microbiome communities are significantly different between R and NR patients

(A) Bar plots of the phylum (left) and genus (right) taxonomic levels in R and NR patients with lung cancer. Relative abundance is used.

(B) Phylum differences between R and NR patients.

(C) LDA score of features with different abundances between R and NR groups. The criteria for differential feature is an LDA score >2.5.

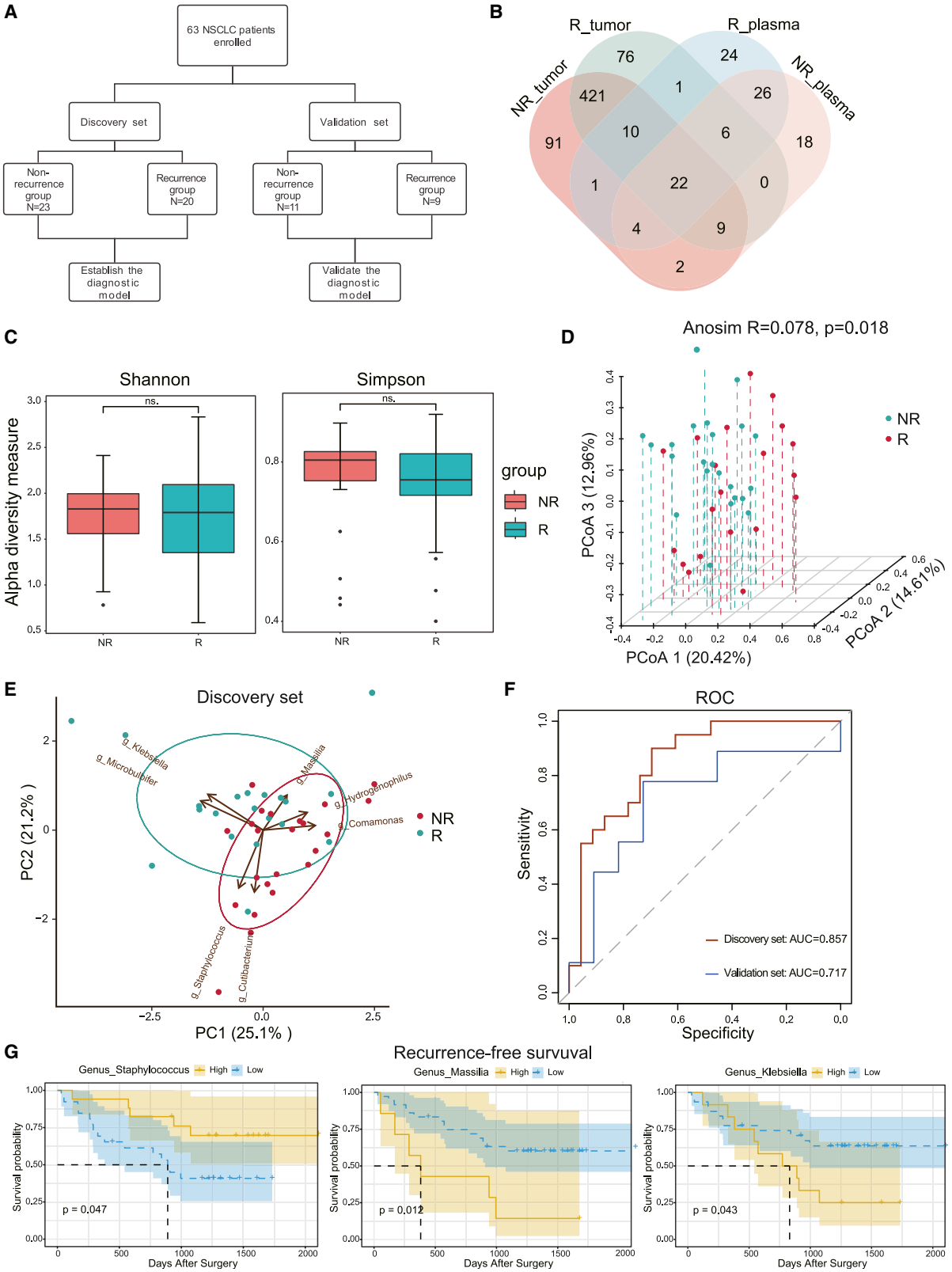
(D) Kaplan-Meier estimates for RFS probability of patients with different abundances of intratumor microbes. Left, *Anaerostipes*, *Helicobacter*, and *Flavonifractor*; right, *Coprococcus*, *Gardnerella*, and *Roseburia*.

(E) Six differentially abundant genera in genus predicting score.

(F) Kaplan-Meier plot of patients with lung cancer defined by genus predicting score.

(G) ROC analysis of genus predicting score as predictive of RFS.

R, recurrence group; NR, non-recurrence group; T, tumor tissues; N, normal lung tissues; LDA, linear discriminant analysis; RFS, recurrence-free survival; ROC, receiver operating characteristics. The error bars indicate the standard deviations. See also Figure S2.



(legend on next page)

different roles in the intestinal tract.[10] The butyrate concentration is relatively lower in other organs and may promote tumor cell proliferation in prostate cancer.[11] We explored the effect of butyrate concentration on the proliferation of lung cancer cells. CCK8 assay and colony-formation assay revealed that butyrate may promote lung cancer cell proliferation at low concentrations (1 mM), while a high concentration (5 mM) of butyrate suppressed the proliferation of lung cancer cells (Figures 4B, 4C, and S4B). Thus, we focused on the effect of butyrate at low concentrations on the migration and invasion of lung cancer cells. We found that butyrate could promote the migration and invasion ability of lung cancer cells (Figures 4D, 4E, S4C, and S4D). Using lung-adenocarcinoma patient-derived organoids (PDO), we found that adding butyrate could promote the production of invasive strands in organoids, indicating that butyrate could promote the invasion ability of lung cancer organoids (Figure 4F).

To further evaluate the function of butyrate *in vivo*, we generated an orthotopic lung cancer model in C57BL/6J mice through the injection of Lewis lung carcinoma (LLC) cells pre-treated with or without butyrate (1 mM). A significant increase in the volume of lung *in situ* tumors was observed in the butyrate group compared to that in the negative control group (Figures 4G and S4E). We then established an LLC tail vein metastasis model in C57BL/6 mice. Mice were aerosolized with vancomycin (10 mg/mL) and neomycin (20 mg/mL) for a week to reduce the interference of commensal lung microbiome. Then, mice were injected intravenously through the tail vein with LLC cells and aerosolized with saline or butyrate (40 or 200 mM). The number of lung metastases in the low-concentration butyrate group (40 mM) was higher than that in the saline group, although the difference was not statistically significant (Figure 4H). The number of lung metastases in the high-concentration butyrate group (200 mM) was similar to that of the saline group. In addition, we found that the maximum volume of lung metastases was significantly larger in the low-concentration butyrate group (40 mM) than that in the saline group (Figure S4F). These findings suggested that butyrate at low concentrations may promote the migration, invasion, and metastasis of lung cancer cells.

Butyrate promotes the progression of lung cancer though upregulation of H19

To further explore the underlying mechanism of butyrate in lung cancer progression, transcriptome sequencing was performed in A549 cells treated with or without butyrate (1 mM). Differential gene expression analysis revealed 2,680 upregulated genes and 656 downregulated genes in the butyrate group (Figure 5A). Gene Ontology and Kyoto Encyclopedia of Genes and Genomes pathway analysis of upregulated genes suggested that various genes related to cell adhesion and connection, such as neuroactive ligand-receptor interaction, extracellular matrix (ECM)-receptor interaction, and cell adhesion molecules were enriched in the butyrate group (Figure S5A). Then, we generated a gene signature using upregulated genes after butyrate treatment and classified patients with lung cancer in TCGA dataset according to disease-free survival (DFS; < or >3 years). Gene set enrichment analysis was performed, and the gene signature upregulated by butyrate was significantly enriched among patients with DFS <3 years (Figure 5B). We further screened the upregulated genes according to intracellular expression (count > 100) and finally got 438 upregulated genes (Table S8), of which H19 was the most significantly upregulated (fold change [FC] = 36.34, $p < 0.001$, Figure 5A). In A549 and H1299 cell lines, we found that butyrate could significantly increase the expression of H19 (Figures 5C and S5B). H19 is a long non-coding RNA that plays an important role in the progression and metastasis of various tumors including lung cancer.[12–14] Higher expression of H19 was correlated with advanced TNM stage and shorter RFS in TCGA non-small cell lung cancer datasets (Figures S5C and S5D).

Previous studies have reported that histone deacetylase 2 (HDAC2) can inhibit the expression of H19 through H3K27 deacetylation to inhibit cancer metastasis and butyrate could inhibit HDAC2.[15,16] Therefore, we speculated that butyrate may increase the expression of H19 through regulating HDAC2 and H3K27 acetylation. In A549 cells, we found that butyrate could inhibit the expression of HDAC2 at both the mRNA and protein levels (Figures 5D and 5E). Inhibition of HDAC2 through butyrate could promote histone H3K27 acetylation (Figure 5E). Chromatin immunoprecipitation experiments confirmed that HDAC2 can bind to the H19 promoter and butyrate treatment can inhibit this binding (Figure 5F). Inhibition of HDAC2 through butyrate increases the level of H3K27 acetylation at the H19 promoter (Figure 5G). Meanwhile. silencing H19 showed no effect on the expressions of HDAC2 and H3K27 acetylation (Figures S5E–S5G). We then inferred that butyrate might promote lung cancer cell progression via the H19.

Since H19 is a non-coding RNA that does not have a protein product, we thus further explore the downstream target of H19. A lot of evidence has shown that H19 could promote the expression of various matrix metalloproteinases (MMPs).[16,17] Transcriptome sequencing data showed that MMP15 was significantly upregulated in the butyrate group (FC = 5.95, $p < 0.001$, Figure 5A). Additionally, PDO assays also confirmed that butyrate treatment induced more MMP15 secretion (Figure 5H). MMP15 was reported to participate in the metastasis of liver cancer, non-small cell lung cancer, and other tumors.[18,19] Immunohistochemistry

Figure 3. Circulating microbiome DNA could distinguish R and NR patients

(A) Flowchart of circulating microbiome DNA analysis.

(B) Venn plot of shared genus in tumor and plasma.

(C) Alpha diversity of circulating microbiome DNA in R and NR patients (Shannon and Simpson indices).

(D) PCoA of circulating microbiome DNA in R and NR patients using Bray-Curtis metric distances of beta diversity.

(E) PCA using circulating microbiome DNA biomarkers in discovery set.

(F) ROC analysis of circulating microbiome DNA signature as predictive of R patients in discovery and validation sets.

(G) Kaplan-Meier estimates for RFS probability based on the abundance levels of microbes in plasma. Left, *Staphylococcus*; middle, *Massilia*; right, *Klebsiella*. R, recurrence group; NR, non-recurrence group; PCoA, principal-coordinate analysis; PCA, principal-component analysis, ROC, receiver operating characteristics. The error bars indicate the standard deviations. See also Figure S3.

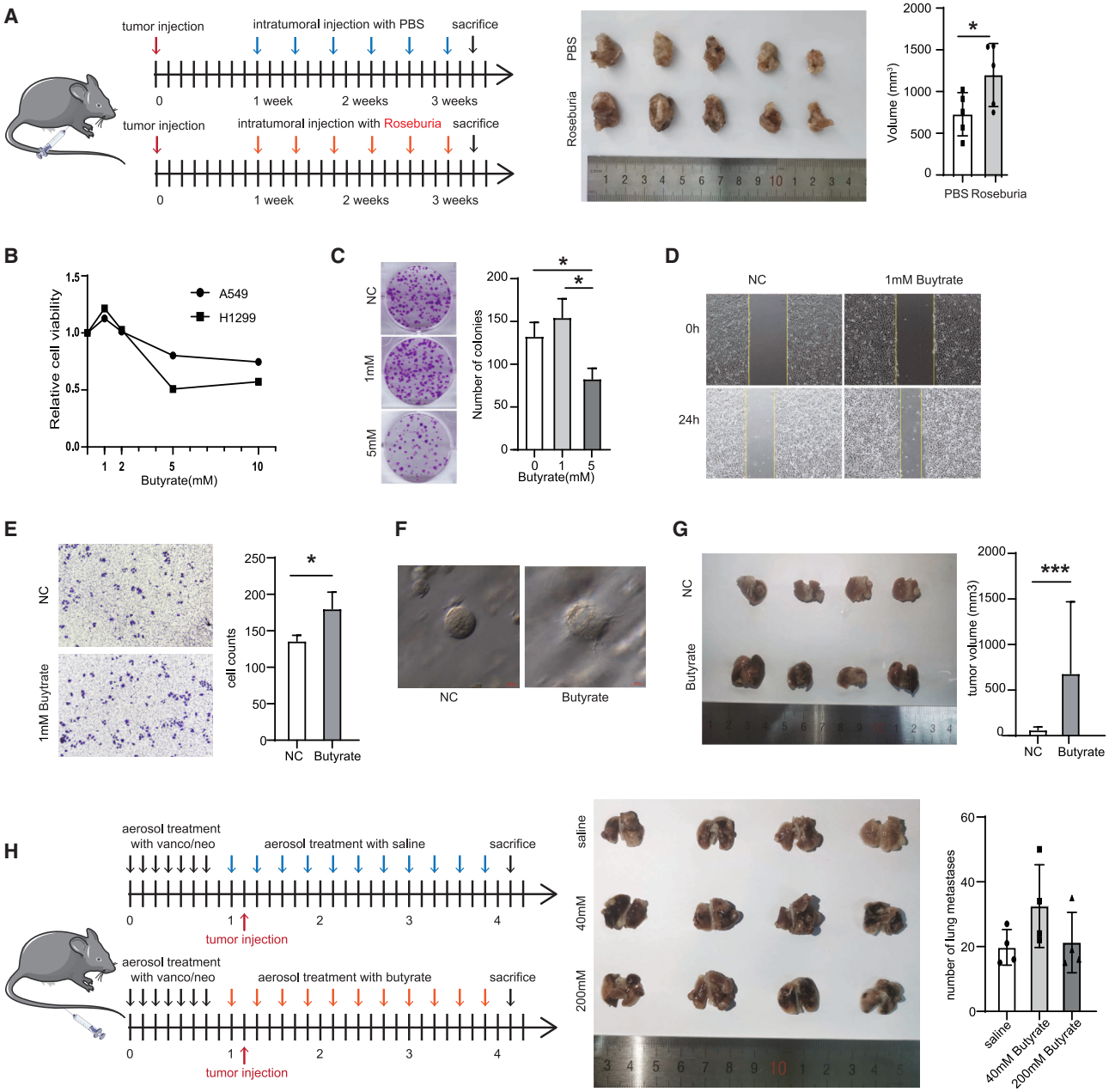


Figure 4. Butyrate promotes the progression of lung cancer at low concentrations

(A) Intratumor injection of *Roseburia* promotes lung cancer growth in the subcutaneous xenograft model.

(B) CCK8 assay of lung cancer cells treated with butyrate at different concentrations. Data depict one representative experiment of three independent experiments; duplicate conditions for each experiment.

(C) Colony-forming assay of butyrate at different concentrations in A549 cells. Data depict one representative experiment of three independent experiments; duplicate conditions for each experiment.

(D) Wound-healing assay of lung cancer A549 cells treated with butyrate (1 mM). Data depict one representative experiment of three independent experiments; duplicate conditions for each experiment.

(E) Transwell assay of lung cancer A549 cells treated with butyrate (1 mM). Data depict one representative experiment of three independent experiments; duplicate conditions for each experiment.

(F) Butyrate treatment induces invasive morphology of lung-adenocarcinoma-derived organoids. Data depict one representative experiment of two independent experiments.

(G) Orthotopic lung cancer model in C57BL/6J mice through injection of LLC cells pre-treated with or without butyrate (1 mM).

(H) LLC tail vein metastasis model in C57BL/6 mice. Mice were aerosolized with vancomycin (10 mg/mL) and neomycin (20 mg/mL) to reduce the interference of the lung microbiome. Then, mice were injected intravenously through the tail vein with LLC cells and aerosolized with saline or butyrate (40 or 200 mM).

p values were calculated by non-paired Student's tests. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$. The error bars indicate the standard deviations. See also Figure S4.

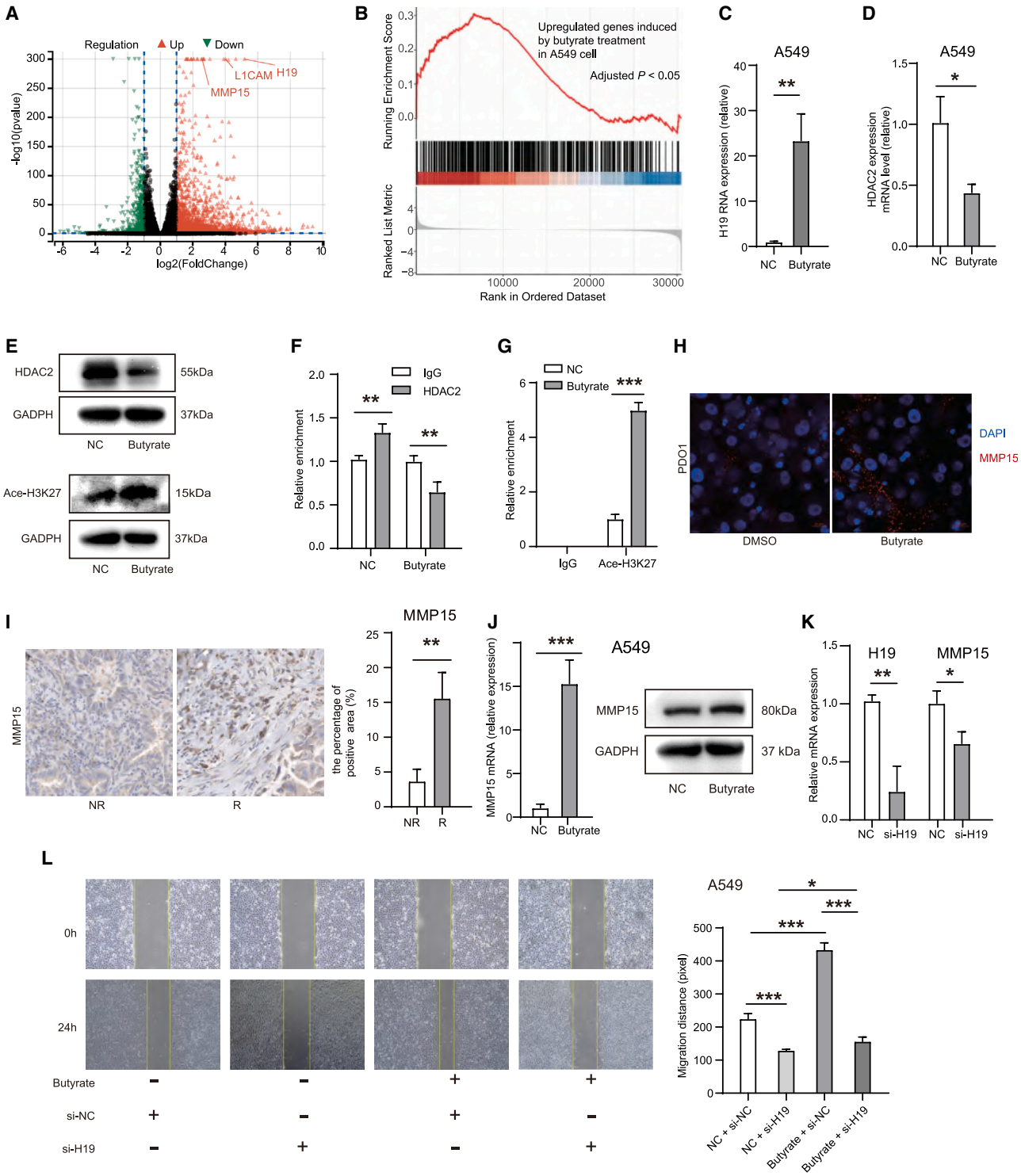


Figure 5. Butyrate promotes lung cancer progression through increasing H19 expression

(A) Volcano plot of differentially expressed genes in A549 cells treated with butyrate (1 mM).

(B) Gene set enrichment analysis of upregulated genes induced by butyrate in TCGA patients with lung cancer.

(C) Expression of H19 in lung cancer cells treated with butyrate (1 mM). Data depict one representative experiment of three independent experiments; duplicate conditions for each experiment.

(D) Expression of HDAC2 in A549 cells treated with butyrate (1 mM).

(E) Detection of HDAC2 (top) and H3K27 acetylation (bottom) by western blot. Data depict one representative experiment of three independent experiments.

(legend continued on next page)

showed that the expression of MMP15 was significantly higher in the tumor tissues of the R group (Figure 5I). We also found that higher expression of MMP15 was correlated with late TNM stage and short RFS in non-small cell lung cancer through TCGA datasets (Figures S5H and S5I). In A549 or H1299 cells, butyrate could significantly increase the expression of MMP15 at both the mRNA and protein levels (Figures 5J and S5J). Conversely, silencing H19 inhibited MMP15 expression in A549 and H1299 cells (Figures 5K and S5K).

We designed functional rescue experiments in which we treated A549 or H1299 cells with low-concentration butyrate and then silenced H19 in cells (Figures 5L and S5L). As shown by wound-healing assay, low-concentration butyrate promoted the migration of lung cancer cells, and the migration advantage conferred by butyrate treatment was partially reversed by H19 silencing. These results suggested that butyrate at low concentrations may inhibit HDAC2 and its binding to the H19 promoter region, thereby increasing the level of H3K27 acetylation at the H19 promoter region, and further increase the expression of H19 and promote lung cancer cell progression.

Butyrate facilitates polarization of M2 macrophages

It has been reported that butyrate could modulate the differentiation of immune cells, such as regulatory T cells and macrophages.[20,21] While M2 macrophages in the tumor microenvironment could promote the invasion and metastasis of tumor cells,[22,23] we hypothesized that low-concentration butyrate may facilitate the polarization of macrophages from M0 to M2, thus promoting tumor metastasis. Bone marrow-derived macrophage (BMDM) or RAW264.7 cells were polarized to the M2 phenotype by interleukin-4 (IL-4) treatment (the M2 group represents macrophages treated with IL-4 in the following experiments), and then M2s were treated with various concentrations of butyrate (0, 0.5, and 1 mM; Figures 6A and S6A). IL-4 successfully induced M2 polarization, which is evidenced by increased expression of *Arg1*, *Ym1*, and *Fizz1* in BMDM or RAW264.7 cells (Figures 6A and S6A). In BMDM cells, the relative expression of *Arg1*, *Fizz1*, and *Ym1* was considerably higher in M2-BMDMs exposed to 0.5 mM butyrate for 24 h (Figure 6A). In RAW264.7 cells, the addition of 0.5 mM butyrate further increased the expression of *Arg1* and *Ym1* relative to the M2 group (Figure S6A). The effect on *Arg1* was most striking in both BMDM or RAW264.7 cells. Western blot analysis confirmed enhanced *Arg1* protein expression after butyrate treatment (Figures 6B and S6B). In a tail vein metastasis model, immunohistochemistry revealed that expression of *Arg1* in the lung metastases of the low-concentration butyrate group was significantly higher than that of the other two groups (Figure 6C). Co-staining of *Arg1* and CD206 also indicated more M2 macrophages in butyrate group in tail vein metastasis model and orthotopic lung cancer model (Figure S6C). Furthermore, co-staining of ARG1 and CD206 also revealed more M2 macrophage presence in tumor tissues of R patients (Figure 6D). Collectively, these data indicated that butyrate facilitated M2 macrophage polarization.

It has been demonstrated that M2 macrophages can promote the migration and invasion of tumor cells by secretion of various cytokines, such as IL-10 and IL-13.[24,25] Culture medium (CM) supernatant of M2 macrophages with or without butyrate treatment was collected and analyzed with the Quantibody mouse cytokine array. The results showed that CM from butyrate-treated M2 macrophages contained more IL-10, IL-13, and MIP3a (Figure 6E), which were reported to correlate with tumor metastasis. Accordingly, the migration- and invasion-promoting effect of CM from M2-BMDMs was significantly increased when M2-BMDMs had been treated with butyrate (Figures 6F, 6G, and S6D).

To further confirm whether macrophages are involved in the butyrate-induced metastasis *in vivo*, we established an LLC tail vein metastasis model and depleted macrophages by clodronate liposomes (Figure 6H). Compared to control liposomes, clodronate liposome treatment remarkably depleted macrophages in peripheral blood (Figure 6I) and liver (Figure S6E), and removal of macrophages attenuated the pro-metastasis function of butyrate (Figures 6J and S6F). Together, these data demonstrated that macrophages are involved the metastasis-promoting function of butyrate.

DISCUSSION

In this study, we have performed a comprehensive analysis of the intratumor microbiome in a cohort of R and NR patients with lung cancer. Overall, we demonstrate that the diversity and composition of tumor microbiome were related to recurrence in lung cancer. Increasing evidence suggests that reduced tumor microbiome diversity is associated with poor survival of patients with cancer.[26,27] We found that lower tumor microbiome diversity was significantly related to reduced RFS. Furthermore, genus predicting scores with specific bacterial genus in tumor or normal lung tissues showed good performance in predicting RFS. cmDNA signatures also showed remarkable predictive performance in discovery and validation sets. Another finding was that butyrate-producing bacteria such as *Roseburia* were disturbed and correlated with reduced RFS. Butyrate was reported to inhibit cell proliferation at high concentrations and promote cell proliferation at low concentrations.[10,11] Interestingly, we demonstrated that

(F) Chromatin immunoprecipitation (ChIP) analysis of HDAC2 enrichment at the promoter of H19 gene in A549 cells treated with or without butyrate (1 mM). Data depict one representative experiment of three independent experiments; duplicate conditions for each experiment.
(G) ChIP analysis of H3K27 acetylation at the promoter of H19 gene in A549 cells treated with or without butyrate (1 mM). Data depict one representative experiment of three independent experiments; duplicate conditions for each experiment.
(H) Butyrate treatment induced more secretion of MMP15 in lung cancer organoids.
(I) Immunohistochemical experiments found MMP15 expression was higher in R patients.
(J) Expression of MMP15 in lung cancer cells treated with butyrate (1 mM). Data depict one representative experiment of three independent experiments.
(K) Expressions of H19 and MMP15 mRNA after transfection of si-H19. Data depict one representative experiment of three independent experiments.
(L) Rescue experiments in A549 cells.
p values were calculated by non-paired Student's tests. Data depict one representative experiment of three independent experiments; duplicate conditions for each experiment. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$. The error bars indicate the standard deviations. See also Figure S5.

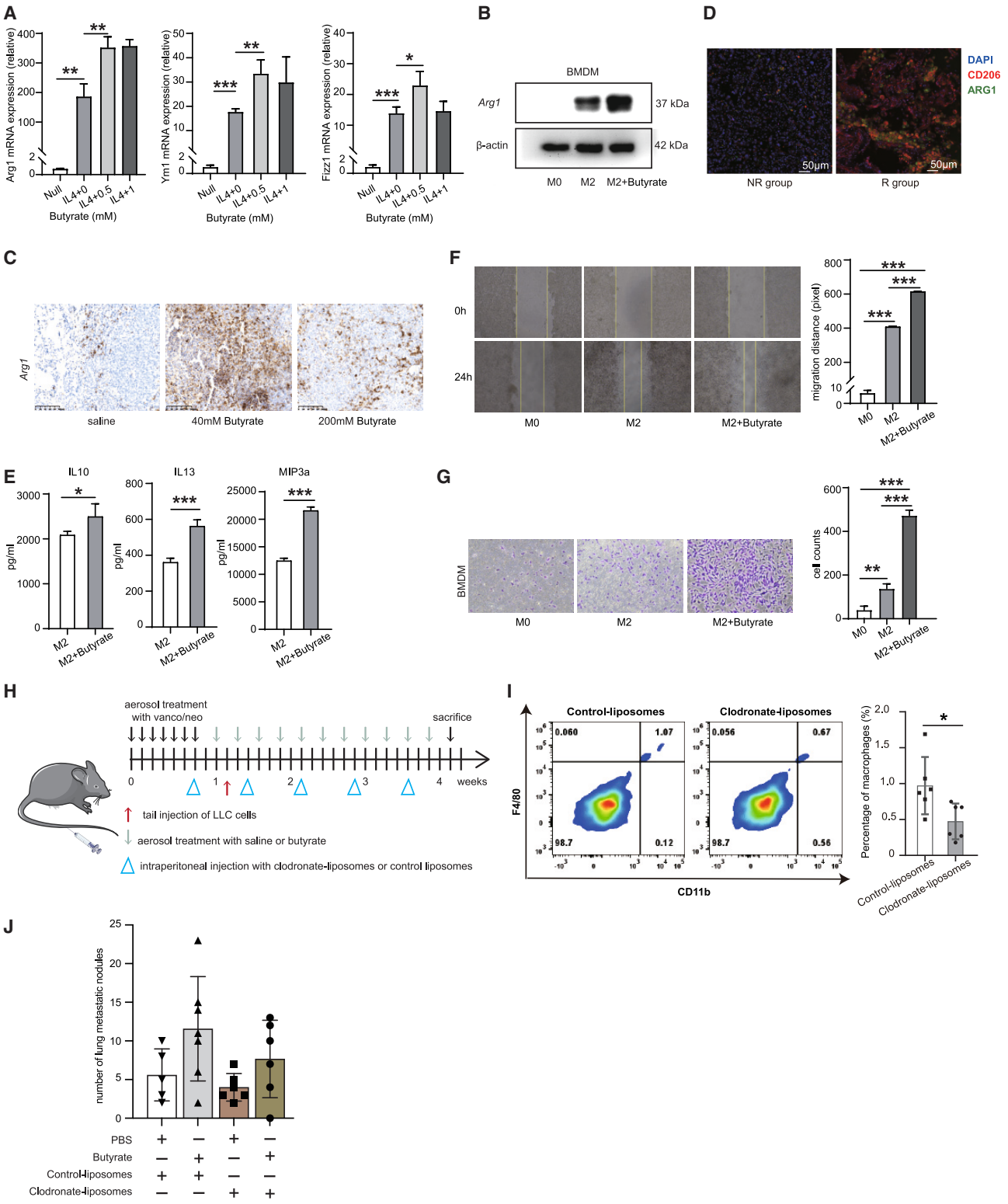


Figure 6. Butyrate facilitates M2 macrophage polarization and invasion-promoting function

(A) BMDMs were treated with different concentrations of butyrate for 24 h. Expression of *Arg1*, *Ym1*, and *Fizz1* was measured by RT-qPCR. Data depict the aggregate of three independent experiments.

(B) Detection of *Arg1* by western blot (0.5 mM butyrate was used). Data depict one representative experiment of three independent experiments.

(legend continued on next page)

butyrate at low concentrations could promote lung cancer progression by upregulating H19 and facilitating M2 macrophage polarization (Figure 7).

Previous studies have shown that lower airway microbiota was associated with the recurrence of lung cancer and could promote tumor progression.[28,29] However, the association between the tumor microbiome and recurrence in lung cancer remains unclear. Our findings suggested that lung cancer tumor microbiome diversity was significantly lower in the R group and related to reduced RFS, indicating that the tumor microbiome may influence tumor recurrence. Importantly, we found a signature of six tumor bacterial genera (*Roseburia*, *Helicobacter*, *Gardnerella*, *Flavonifractor*, *Coprococcus*, and *Anaerostipes*), which may serve as potential biomarkers to stratify patients for recurrence. Furthermore, Poore et al. proposed a novel cancer diagnostic approach with preferable accuracy through microbiome analyses of blood.[30] They found that cmDNA could diagnosis lung cancer. In this study, we found that cmDNA signatures showed good performance for recurrence prediction in both discovery and validation sets. cmDNA has the potential to become a non-invasive biomarker for recurrence prediction. However, large-scale studies are needed to further validate the reliability of tumor or cmDNA signatures in cancer diagnoses.

Roseburia is a well-known butyrate-producing bacteria and was reported to inhibit inflammation in the intestinal tract.[31] However, *Roseburia* was enriched in both tumor and normal tissues of the R group and was related to reduced RFS in the current study. Similar to our findings, Peters et al. found that *Roseburia* in normal lung tissues was associated with reduced survival in patients with lung cancer.[32] In a subcutaneous xenograft mouse model, we found that intratumoral injection of *Roseburia* increased tumor volume. These findings indicated that microbiota may perform different functions outside the gut. Previous studies have demonstrated that butyrate may play different roles in the intestinal tract according to its concentration.[10] The butyrate concentration is relative lower in other organs and may promote tumor cell proliferation in prostate cancer.[11] We found that various butyrate-producing bacteria such as *Roseburia* were disturbed in the tumor and normal lung tissue of patients with lung cancer with recurrence. Functional assays showed that butyrate at low concentrations could promote lung cancer cell proliferation to a certain extent, while butyrate at high concentrations inhibits lung cancer cell proliferation. Moreover, butyrate at low concentrations could promote lung cancer cell migration and invasion. Mechanistically, genes upregulated by low-concentration butyrate enriched pathways related to cell adhesion and connection, such as ECM-receptor interaction, and cell adhesion molecules. Cell adhesion and connection molecules play important roles in tumor progression and metastasis.[33] H19 is a long non-coding RNA that plays an important role in the progression and metastasis of various tumors including lung cancer.[12–14] We found that butyrate treatment significantly increased the expression of H19. Acetylation is the most common post-translational modification of histones and is regulated by HDAC.[34,35] Histone H3K27 acetylation is a hallmark of transcriptional activation of gene expression.[36] Previous studies have found that histone H3K27 acetylation promotes cancer progression in colon cancer, lung cancer, breast cancer, and other tumors.[36–38] We found that butyrate could inhibit HDAC2 expression and increase histone H3K27 acetylation. Hu et al. demonstrated that HDAC2 could inhibit the expression of H19 through H3K27 deacetylation to inhibit cancer metastasis.[16] Our findings showed that butyrate at low concentrations could inhibit HDAC2 and its binding to the H19 promoter region, thereby increasing the level of H3K27 acetylation in the H19 promoter region, and further increase the expression of H19 and promote lung cancer cell progression. In addition, we also found that MMP15, which was reported to promote lung cancer metastasis, was upregulated by butyrate. H19 could regulate multiple MMP expression through sponging microRNAs.[16,17] We found that silencing H19 could inhibit MMP15 expression. However, the mechanism of H19 regulating MMP15 needs further exploration.

In the tumor microenvironment, M2 macrophages play important roles in tumor progression and metastasis.[22,23] Ji et al. found that butyrate facilitates M2 macrophage polarization in BMDMs.[39] In the current study, we found that butyrate promotes IL-4-mediated M2 macrophage polarization in RAW264.7 and BMDM cells. The specific marker of the tumor-associated macrophage, *Arg1*, was significantly upregulated by butyrate. M2 macrophages can enhance the migration and invasion ability of tumor cells mainly by secreting cytokines, chemokines, and proteases.[22,23] Notably, using culture supernatant, we found that butyrate could enhance the ability of M2 macrophages to promote tumor cell migration and invasion. These findings indicated that butyrate may promote tumor progression and metastasis by facilitating M2 macrophage polarization and function. Further studies are needed to identify the underlying mechanism through which butyrate regulates M2 macrophage polarization and function.

Conclusion

In conclusion, we found that the intratumor microbiome and cmDNA signatures are promising biomarkers in determining

(C) Expression of *Arg1* in lung metastases in mice aerosolized with saline or butyrate (40 or 200 mM).
(D) Co-staining of ARG1 and CD206 in the tumor tissues of NR and R groups.
(E) Elevated IL-10, IL-13, and MIP3a were found in culture media from M2 macrophages treated with butyrate.
(F) Migration ability of LLC cells treated with culture media from BMDMs treated with or without butyrate (0.5 mM). Data depict one representative experiment of three independent experiments; duplicate conditions for each experiment.
(G) Invasion ability of LLC cells treated with culture medium from macrophages treated with or without butyrate (0.5 mM). Data depict one representative experiment of three independent experiments; duplicate conditions for each experiment.
(H) LLC tail vein metastasis model in C57BL/6 mice with or without removal of the macrophages by clodronate liposomes.
(I) Detection of macrophages in peripheral blood of mice treated with clodronate liposomes or control liposomes. Data depict one representative experiment of two independent experiments.
(J) Removal of macrophages attenuated the pro-metastasis function of butyrate.
p values were calculated by non-paired Student's tests. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$. The error bars indicate the standard deviations. See also Figure S6.

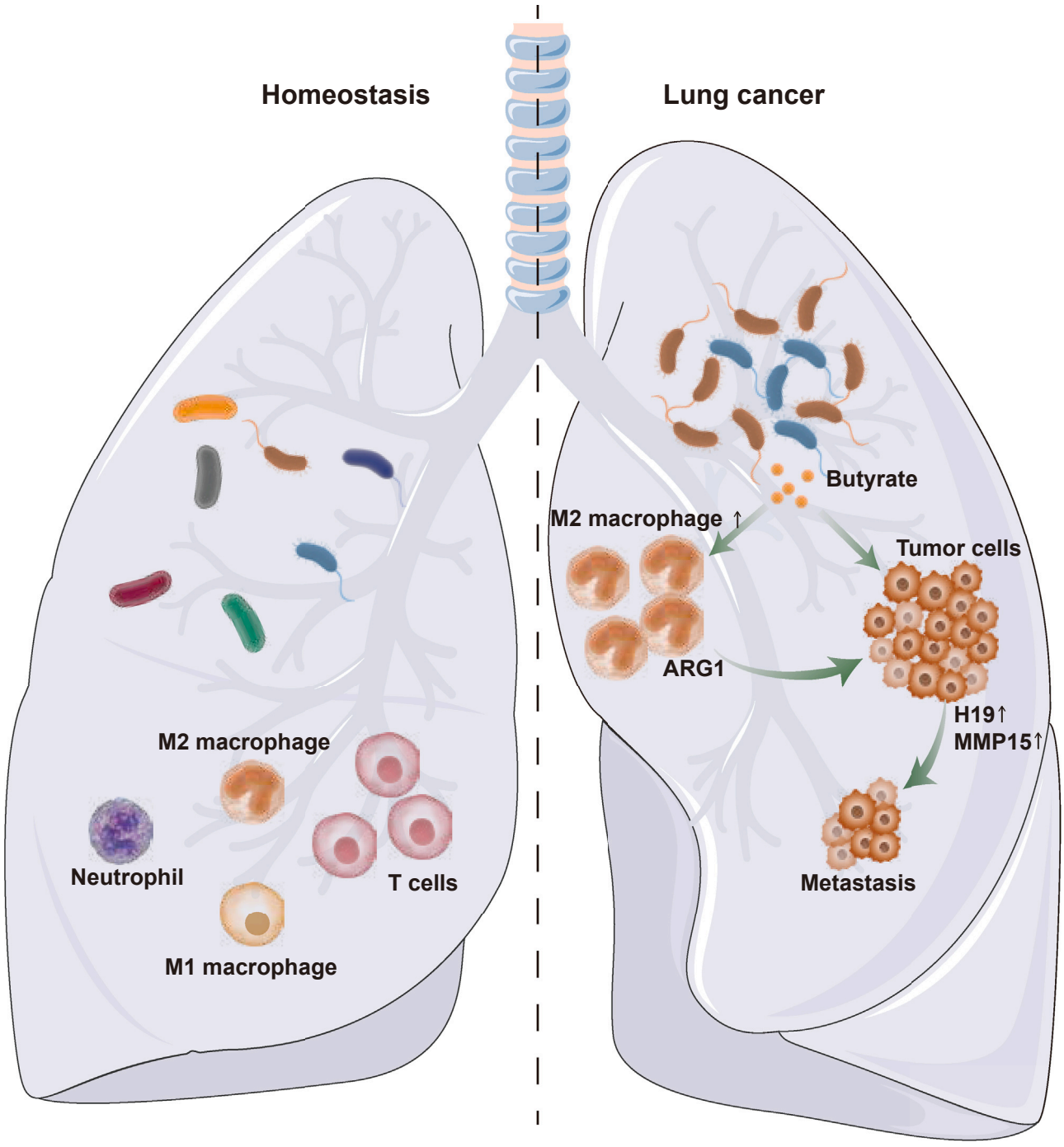


Figure 7. Schematic diagram of the intratumor butyrate-producing bacteria and low-concentration-butyrate promoting lung cancer progression

the recurrence of patients with lung cancer. The dysbiosis of butyrate-producing bacteria may contribute to the tumor recurrence. Butyrate at low concentrations may promote tumor progression and metastasis through upregulating H19 and facilitating M2 macrophage polarization and function.

Limitations of the study

First, sample size of the validation cohorts was limited; thus, caution should be taken when expanding to different populations. Second, the concentration of butyrate may be dynamic in the tumor microenvironment and difficult to be accurately detected. Future analysis using a novel targeted metabolomics

method might provide more evidence. Third, we found that removal of macrophages attenuated the pro-metastasis function of butyrate; however, no statistical difference was achieved because of the limited number of mice in each group. In addition, neutralizing antibodies against the cytokines are needed to further explore the mechanism of butyrate on M2 macrophage polarization.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Patients and samples
  - Cell lines and reagents
  - Bone marrow–derived macrophage (BMDM) isolation and differentiation
  - Lung cancer derived organoids culture
  - Animals
- METHOD DETAILS
  - DNA extraction and 16s gene sequencing
  - Circulating microbiome DNA sequencing and analysis
  - RNA sequencing and analysis
  - RNA extraction and qRT-PCR analysis
  - Construction of small interfering RNAs (siRNAs) and transfection of cells
  - Cell proliferation, migration, and invasion
  - Chromatin immunoprecipitation (ChIP) and qPCR
  - Western blot analysis
  - Histology immunohistochemistry (IHC) and Immunofluorescence (IF)
  - Conditioned medium (CM) preparation
  - Cytokine detection in supernatant of cell culture media
  - Flow cytometry analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xcrm.2024.101488.

ACKNOWLEDGMENTS

We express our thanks to Dr. Chengyang Liu and Jingyi Yang (Institute of Systems Biomedicine, Peking University Health Science Center) for the assistance with animal experiments. This study was supported by the Peking University People's Hospital Scientific Research Development Founds (RZ2022-04 and RDH2020-10), the Peking University Medicine Sailing Program for Young Scholars' Scientific & Technological Innovation (BMU2023YFJHMX010), the National Natural Science Foundation of China (82173386), the Beijing Nova Program (20230484314), the Jiangsu Province Science and Technology Plan Project Key R&D Program (BE2023783), the Open Project of Jiangsu Provincial Science and Technology Resources (Clinical Resources) Coordination Service Platform (TC2022B010), the Key Research Project of Jiangsu Provincial Health Commission (K2023051), and the Shanghai Pulmonary Hospital National Natural Science Cultivate Foundation (fkzr2352).

AUTHOR CONTRIBUTIONS

M.Q., Y. Li, and Y.M. designed the study. Y.M., H.C., H.L., M.Z., X.Z., and W.W. carried out all aspects of the research and experiments. Y.M. and Jun Wang performed the 16S rRNA sequencing data analysis. Y.M., H.C., and Y. Lu performed cmDNA sequencing data analysis. S.W. and Y. Li collected clinical data. Y.M. and H.C. wrote the manuscript. S.W., Jun Wang, Y. Li, Jie Wang, and M.Q. revised the manuscript. M.Q. supervised the project.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 18, 2022
Revised: December 1, 2023
Accepted: March 5, 2024
Published: April 1, 2024

REFERENCES

1. Chen, W., Zheng, R., Baade, P.D., Zhang, S., Zeng, H., Bray, F., Jemal, A., Yu, X.Q., and He, J. (2016). Cancer statistics in China, 2015. CA A Cancer J. Clin. *66*, 115–132. https://doi.org/10.3322/caac.21338.
2. Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., and Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA A Cancer J. Clin. *71*, 209–249. https://doi.org/10.3322/caac.21660.
3. Bagcchi, S. (2017). Lung cancer survival only increases by a small amount despite recent treatment advances. Lancet Respir. Med. *5*, 169. https://doi.org/10.1016/S2213-2600(17)30041-3.
4. van den Berg, L.L., Klinkenberg, T.J., Groen, H.J.M., and Widder, J. (2015). Patterns of Recurrence and Survival after Surgery or Stereotactic Radiotherapy for Early Stage NSCLC. J. Thorac. Oncol. *10*, 826–831. https://doi.org/10.1097/JTO.0000000000000483.
5. Hanahan, D. (2022). Hallmarks of Cancer: New Dimensions. Cancer Discov. *12*, 31–46. https://doi.org/10.1158/2159-8290.CD-21-1059.
6. Nejman, D., Livyatan, I., Fuks, G., Gavert, N., Zwang, Y., Geller, L.T., Rotter-Maskowitz, A., Weiser, R., Mallel, G., Gigi, E., et al. (2020). The human tumor microbiome is composed of tumor type-specific intracellular bacteria. Science *368*, 973–980. https://doi.org/10.1126/science.aay9189.
7. Hezaveh, K., Shinde, R.S., Klötgen, A., Halaby, M.J., Lamorte, S., Ciudad, M.T., Quevedo, R., Neufeld, L., Liu, Z.Q., Jin, R., et al. (2022). Tryptophan-derived microbial metabolites activate the aryl hydrocarbon receptor in tumor-associated macrophages to suppress anti-tumor immunity. Immunity *55*, 324–340.e8. https://doi.org/10.1016/j.immuni.2022.01.006.
8. Ma, Y., Qiu, M., Wang, S., Meng, S., Yang, F., and Jiang, G. (2021). Distinct tumor bacterial microbiome in lung adenocarcinomas manifested as radiological subsolid nodules. Transl. Oncol. *14*, 101050. https://doi.org/10.1016/j.tranon.2021.101050.
9. Chen, H., Ma, Y., Liu, Z., Li, J., Li, X., Yang, F., and Qiu, M. (2021). Circulating microbiome DNA: An emerging paradigm for cancer liquid biopsy. Cancer Lett. *521*, 82–87. https://doi.org/10.1016/j.canlet.2021.08.036.
10. Lupton, J.R. (2004). Microbial degradation products influence colon cancer risk: the butyrate controversy. J. Nutr. *134*, 479–482. https://doi.org/10.1093/jn/134.2.479.
11. Matsushita, M., Fujita, K., Hayashi, T., Kayama, H., Motooka, D., Hase, H., Jingushi, K., Yamamichi, G., Yumiba, S., Tomiyama, E., et al. (2021). Gut Microbiota-Derived Short-Chain Fatty Acids Promote Prostate Cancer Growth via IGF1 Signaling. Cancer Res. *81*, 4014–4026. https://doi.org/10.1158/0008-5472.CAN-20-4090.

12. Singh, N., Ramnarine, V.R., Song, J.H., Pandey, R., Padi, S.K.R., Nouri, M., Olive, V., Kobelev, M., Okumura, K., McCarthy, D., et al. (2021). The long noncoding RNA H19 regulates tumor plasticity in neuroendocrine prostate cancer. Nat. Commun. *12*, 7349. https://doi.org/10.1038/s41467-021-26901-9.

13. Yang, J., Qi, M., Fei, X., Wang, X., and Wang, K. (2021). LncRNA H19: A novel oncogene in multiple cancers. Int. J. Biol. Sci. *17*, 3188–3208. https://doi.org/10.7150/ijbs.62573.

14. Zhang, Y., Huang, W., Yuan, Y., Li, J., Wu, J., Yu, J., He, Y., Wei, Z., and Zhang, C. (2020). Long non-coding RNA H19 promotes colorectal cancer metastasis via binding to hnRNPA2B1. J. Exp. Clin. Cancer Res. *39*, 141. https://doi.org/10.1186/s13046-020-01619-6.

15. Kim, D.S., Kwon, J.E., Lee, S.H., Kim, E.K., Ryu, J.G., Jung, K.A., Choi, J.W., Park, M.J., Moon, Y.M., Park, S.H., et al. (2018). Attenuation of Rheumatoid Inflammation by Sodium Butyrate Through Reciprocal Targeting of HDAC2 in Osteoclasts and HDAC8 in T Cells. Front. Immunol. *9*, 1525. https://doi.org/10.3389/fimmu.2018.01525.

16. Hu, X.T., Xing, W., Zhao, R.S., Tan, Y., Wu, X.F., Ao, L.Q., Li, Z., Yao, M.W., Yuan, M., Guo, W., et al. (2020). HDAC2 inhibits EMT-mediated cancer metastasis by downregulating the long noncoding RNA H19 in colorectal cancer. J. Exp. Clin. Cancer Res. *39*, 270. https://doi.org/10.1186/s13046-020-01783-9.

17. Li, L., Han, T., Liu, K., Lei, C.G., Wang, Z.C., and Shi, G.J. (2019). LncRNA H19 promotes the development of hepatitis B related hepatocellular carcinoma through regulating microRNA-22 via EMT pathway. Eur. Rev. Med. Pharmacol. Sci. *23*, 5392–5401. https://doi.org/10.26355/eurrev_201906_18208.

18. Zheng, S., Wu, H., Wang, F., Lv, J., Lu, J., Fang, Q., Wang, F., Lu, Y., Zhang, S., Xu, Y., et al. (2019). The oncoprotein HBXIP facilitates metastasis of hepatocellular carcinoma cells by activation of MMP15 expression. Cancer Manag. Res. *11*, 4529–4540. https://doi.org/10.2147/CMAR.S198783.

19. Jia, Y.C., Wang, J.Y., Liu, Y.Y., Li, B., Guo, H., and Zang, A.M. (2019). LncRNA MAFG-AS1 facilitates the migration and invasion of NSCLC cell via sponging miR-339-5p from MMP15. Cell Biol. Int. *43*, 384–393. https://doi.org/10.1002/cbin.11092.

20. Yip, W., Hughes, M.R., Li, Y., Cait, A., Hirst, M., Mohn, W.W., and McNagny, K.M. (2021). Butyrate Shapes Immune Cell Fate and Function in Allergic Asthma. Front. Immunol. *12*, 628453. https://doi.org/10.3389/fimmu.2021.628453.

21. Liang, L., Liu, L., Zhou, W., Yang, C., Mai, G., Li, H., and Chen, Y. (2022). Gut microbiota-derived butyrate regulates gut mucus barrier repair by activating the macrophage/WNT/ERK signaling pathway. Clin. Sci. *136*, 291–307. https://doi.org/10.1042/CS20210778.

22. Rhee, I. (2016). Diverse macrophages polarization in tumor microenvironment. Arch Pharm. Res. (Seoul) *39*, 1588–1596. https://doi.org/10.1007/s12272-016-0820-y.

23. Lan, J., Sun, L., Xu, F., Liu, L., Hu, F., Song, D., Hou, Z., Wu, W., Luo, X., Wang, J., et al. (2019). M2 Macrophage-Derived Exosomes Promote Cell Migration and Invasion in Colon Cancer. Cancer Res. *79*, 146–158. https://doi.org/10.1158/0008-5472.CAN-18-0014.

24. Zhang, H., Li, R., Cao, Y., Gu, Y., Lin, C., Liu, X., Lv, K., He, X., Fang, H., Jin, K., et al. (2022). Poor Clinical Outcomes and Immunoevasive Contexture in Intratumoral IL-10-Producing Macrophages Enriched Gastric Cancer Patients. Ann. Surg. *275*, e626–e635. https://doi.org/10.1097/SLA.0000000000004037.

25. Raggi, C., Correnti, M., Sica, A., Andersen, J.B., Cardinale, V., Alvaro, D., Chiorino, G., Forti, E., Glaser, S., Alpini, G., et al. (2017). Cholangiocarcinoma stem-like subset shapes tumor-initiating niche by educating associated macrophages. J. Hepatol. *66*, 102–115. https://doi.org/10.1016/j.jhep.2016.08.012.

26. Riquelme, E., Zhang, Y., Zhang, L., Montiel, M., Zoltan, M., Dong, W., Quesada, P., Sahin, I., Chandra, V., San Lucas, A., et al. (2019). Tumor Microbiome Diversity and Composition Influence Pancreatic Cancer Outcomes. Cell *178*, 795–806.e12. https://doi.org/10.1016/j.cell.2019.07.008.

27. Boesch, M., Baty, F., Albrich, W.C., Flatz, L., Rodriguez, R., Rothschild, S.I., Joerger, M., Früh, M., and Brutsche, M.H. (2021). Local tumor microbial signatures and response to checkpoint blockade in non-small cell lung cancer. OncoImmunology *10*, 1988403. https://doi.org/10.1080/2162402X.2021.1988403.

28. Patnaik, S.K., Cortes, E.G., Kannisto, E.D., Punnanitinont, A., Dhillon, S.S., Liu, S., and Yendamuri, S. (2021). Lower airway bacterial microbiome may influence recurrence after resection of early-stage non-small cell lung cancer. J. Thorac. Cardiovasc. Surg. *161*, 419–429.e16. https://doi.org/10.1016/j.jtcvs.2020.01.104.

29. Tsay, J.C.J., Wu, B.G., Sulaiman, I., Gershner, K., Schluger, R., Li, Y., Yie, T.A., Meyn, P., Olsen, E., Perez, L., et al. (2021). Lower Airway Dysbiosis Affects Lung Cancer Progression. Cancer Discov. *11*, 293–307. https://doi.org/10.1158/2159-8290.CD-20-0263.

30. Poore, G.D., Kopylova, E., Zhu, Q., Carpenter, C., Fraraccio, S., Wandro, S., Kosciolek, T., Janssen, S., Metcalf, J., Song, S.J., et al. (2020). Microbiome analyses of blood and tissues suggest cancer diagnostic approach. Nature *579*, 567–574. https://doi.org/10.1038/s41586-020-2095-1.

31. Shen, Z., Zhu, C., Quan, Y., Yang, J., Yuan, W., Yang, Z., Wu, S., Luo, W., Tan, B., and Wang, X. (2018). Insights into Roseburia intestinalis which alleviates experimental colitis pathology by inducing anti-inflammatory responses. J. Gastroenterol. Hepatol. *33*, 1751–1760. https://doi.org/10.1111/jgh.14144.

32. Peters, B.A., Hayes, R.B., Goparaju, C., Reid, C., Pass, H.I., and Ahn, J. (2019). The Microbiome in Lung Cancer Tissue and Recurrence-Free Survival. Cancer Epidemiol. Biomarkers Prev. *28*, 731–740. https://doi.org/10.1158/1055-9965.EPI-18-0966.

33. Läubli, H., and Borsig, L. (2019). Altered Cell Adhesion and Glycosylation Promote Cancer Immune Suppression and Metastasis. Front. Immunol. *10*, 2120. https://doi.org/10.3389/fimmu.2019.02120.

34. Seto, E., and Yoshida, M. (2014). Erasers of histone acetylation: the histone deacetylase enzymes. Cold Spring Harbor Perspect. Biol. *6*, a018713. https://doi.org/10.1101/cshperspect.a018713.

35. Verdone, L., Caserta, M., and Di Mauro, E. (2005). Role of histone acetylation in the control of gene expression. Biochem. Cell. Biol. *83*, 344–353. https://doi.org/10.1139/o05-041.

36. Ye, P., Lv, X., Aizemaiti, R., Cheng, J., Xia, P., and Di, M. (2020). H3K27ac-activated LINC00519 promotes lung squamous cell carcinoma progression by targeting miR-450b-5p/miR-515-5p/YAP1 axis. Cell Prolif. *53*, e12797. https://doi.org/10.1111/cpr.12797.

37. Liu, D., Zhang, H., Cong, J., Cui, M., Ma, M., Zhang, F., Sun, H., and Chen, C. (2020). H3K27 acetylation-induced lncRNA EIF3J-AS1 improved proliferation and impeded apoptosis of colorectal cancer through miR-3163/YAP1 axis. J. Cell. Biochem. *121*, 1923–1933. https://doi.org/10.1002/jcb.29427.

38. Han, M., Qian, X., Cao, H., Wang, F., Li, X., Han, N., Yang, X., Yang, Y., Dou, D., Hu, J., et al. (2020). lncRNA ZNF649-AS1 Induces Trastuzumab Resistance by Promoting ATG5 Expression and Autophagy. Mol. Ther. *28*, 2488–2502. https://doi.org/10.1016/j.ymthe.2020.07.019.

39. Ji, J., Shu, D., Zheng, M., Wang, J., Luo, C., Wang, Y., Guo, F., Zou, X., Lv, X., Li, Y., et al. (2016). Microbial metabolite butyrate facilitates M2 macrophage polarization and function. Sci. Rep. *6*, 24838. https://doi.org/10.1038/srep24838.

40. Mallick, H., Rahnavard, A., McIver, L.J., Ma, S., Zhang, Y., Nguyen, L.H., Tickle, T.L., Weingart, G., Ren, B., Schwager, E.H., et al. (2021). Multivariable association discovery in population-scale meta-omics studies. PLoS Comput. Biol. *17*, e1009442. https://doi.org/10.1371/journal.pcbi.1009442.

41. Breiman, L. (2001). Random forests. Mach Learn *45*, 5–32.

42. Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.C., and Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinf. *12*, 77. https://doi.org/10.1186/1471-2105-12-77.

43. Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169. https://doi.org/10.1093/bioinformatics/btu638.

44. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.

45. Duncan, S.H., Hold, G.L., Barcenilla, A., Stewart, C.S., and Flint, H.J. (2002). Roseburia intestinalis sp. nov., a novel saccharolytic, butyrate-producing bacterium from human faeces. Int. J. Syst. Evol. Microbiol. *52*, 1615–1620. https://doi.org/10.1099/00207713-52-5-1615.

46. Cao, J., Dong, R., Jiang, L., Gong, Y., Yuan, M., You, J., Meng, W., Chen, Z., Zhang, N., Weng, Q., et al. (2019). LncRNA-MM2P Identified as a Modulator of Macrophage M2 Polarization. Cancer Immunol. Res. *7*, 292–305. https://doi.org/10.1158/2326-6066.CIR-18-0145.

47. Doki, Y., Murakami, K., Yamaura, T., Sugiyama, S., Misaki, T., and Saiki, I. (1999). Mediastinal lymph node metastasis model by orthotopic intrapulmonary implantation of Lewis lung carcinoma cells in mice. Br. J. Cancer *79*, 1121–1126. https://doi.org/10.1038/sj.bjc.6690178.

48. Magoč, T., and Salzberg, S.L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics *27*, 2957–2963. https://doi.org/10.1093/bioinformatics/btr507.

49. Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl. Environ. Microbiol. *73*, 5261–5267. https://doi.org/10.1128/AEM.00062-07.

50. Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S., and Huttenhower, C. (2011). Metagenomic biomarker discovery and explanation. Genome Biol. *12*, R60. https://doi.org/10.1186/gb-2011-12-6-r60.

51. Qiu, M., Xu, Y., Yang, X., Wang, J., Hu, J., Xu, L., and Yin, R. (2014). CCAT2 is a lung adenocarcinoma-specific long non-coding RNA and promotes invasion of non-small cell lung cancer. Tumour Biol. *35*, 5375–5380. https://doi.org/10.1007/s13277-014-1700-z.

52. Schacht, V., and Kern, J.S. (2015). Basics of immunohistochemistry. J. Invest. Dermatol. *135*, 1–4. https://doi.org/10.1038/jid.2014.541.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit anti-Beta Actin	Proteintech	Cat# 20536-1-AP; RRID: AB_10700003
Rabbit anti-GAPDH	Proteintech	Cat# 10494-1-AP; RRID: AB_2263076
Rabbit IgG Control Polyclonal	Proteintech	Cat# 30000-0-AP; RRID: AB_2819035
Rabbit anti-Arg1	Proteintech	Cat# 16001-1-AP; RRID: AB_2289842
Rabbit anti-HDAC2	Proteintech	Cat# 12922-3-AP; RRID: AB_2118516
Rabbit anti-MMP15	Immunoway	Cat# YT2911; RRID: AB_3095584
Rabbit anti-Histone H3 (acetyl 27)	Abcam	Cat# ab177178; RRID: AB_2828007
Rabbit anti-mouse/human Arg1	Servicebio	Cat# GB11285; RRID: AB_3095585
Mouse anti-human CD206	Proteintech	Cat# 60143-1-Ig; RRID: AB_2144924
Rabbit anti-mouse/human CD206	Proteintech	Cat# 18704-1-AP; RRID: AB_10597232
Rabbit anti-F4/80	CST	Cat# 70076S; RRID: AB_2799771
FITC CD11b	BD Biosciences	Cat# 561691; RRID: AB_10893590
PE Rat anti-mouse F4/80	BD Biosciences	Cat# 563899; RRID: AB_2738475
Bacterial and virus strains		
Roseburia intestinalis	DSMZ	14610
Biological samples		
Lung tumor and normal tissues, plasma from patients	specimen repository in Peking University People's Hospital and Jiangsu Cancer Hospital	N/A
Chemicals, peptides, and recombinant proteins		
DMEM	Gibico	2125039
RPMI	Gibico	2122752
HBSS	Lonza	04-315Q
Fetal Bovine Serum	Gibico	26010074
Penicillin-Streptomycin	Gibico	15140122
Sodium butyrate	Solarbio	IS0190
Recombinant Murine M-CSF	Peprotech	315–02
Recombinant Murine IL-4	Peprotech	214–14
Human EGF	Peprotech	AF-100-15
Human FGF10	Peprotech	AF-100-26
Human FGF4	Peprotech	AF-100-31
B27	Gibico	17504–044
DNase	Sigma-Aldrich	AMPD1
Collagenase/dispase	Roche	COLLDISP-RO
Matrigel	Corning	356231
Lipofectamine® RNAiMAX	Invitrogen	13778100
RIPA buffer	Cell Signaling Technology	9806S
protease inhibitor and phosphatase inhibitor cocktail	Sigma-Aldrich	MSSAFE-1VL
Critical commercial assays		
Qubit dsDNA HS Assay Kit	Thermo Fisher	Q32851
VAHTS® Universal DNA Library Prep Kit	Vazyme	ND607-01
NEBNext® Ultra™ RNA Library Prep Kit	NEB	E7530L
Cell Counting Kit-8	Biosharp	BS350A

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
ChIP assay kit	Beyotime Institute of Biotechnology	P2078
Quantibody® Mouse TH17 Array 1	RayBiotech	QAM-TH17-1
Deposited data		
The raw 16S rRNA sequencing data	This paper	BIG: HRA006107
Experimental models: Cell lines		
A549	the Cell Bank of the Chinese Academy of Sciences	SCSP-503
NCI-H1299	the Cell Bank of the Chinese Academy of Sciences	TCHu160
LLC	the Cell Bank of the Chinese Academy of Sciences	TCM 7
RAW264.7	the Cell Bank of the Chinese Academy of Sciences	TCM13
Experimental models: Organisms/strains		
Mouse: C57BL/6	Charles River	N/A
Oligonucleotides		
See Table S9 for primer sequences		N/A
Software and algorithms		
FLASH software (version 1.2.11)	https://ccb.jhu.edu/software/FLASH/index.shtml	N/A
FASTX Toolkit (version 1.2.11)	http://hannonlab.cshl.edu/fastx_toolkit/	N/A
Bowtie2 (v2.3.5.1)	https://bowtie-bio.sourceforge.net/bowtie2/index.shtml	N/A
MaAslin2 software	Mallick et al.[40]	N/A
randomForest	Breiman et al.[41]	N/A
pROC	Robin et al.[42]	N/A
HISAT2 v2.1.0	https://daehwankimlab.github.io/hisat2/	N/A
HTSeq v0.6.0	Anders et al.[43]	N/A
DESeq2	Love et al.[44]	N/A
GSEABase	https://bioconductor.org/packages/release/bioc/html/GSEABase.html	N/A
R version 3.6.1	R Development Core Team	https://www.r-project.org
GraphPad Prism 8	GraphPad	www.graphpad.com/scientificsoftware/prism
SPSS Version 22	IBM Corporation	N/A
ImageJ	https://imagej.nih.gov/ij	N/A
FlowJo	BD Biosciences	https://www.flowjo.com/solutions/flowjo
Custom code in this paper	Mendeley Data	https://doi.org/10.17632/j8j49wk8bw.1
Other		
H19 and MMP15 mRNA expression and DFS in TCGA-NSCLC samples	Gene Expression Profiling Interactive Analysis 2(GEPIA2)	http://gepia2.cancer-pku.cn/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Mantang Qiu (qiumantang@163.com).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The raw 16S rRNA sequencing data reported in this paper have been deposited in the Genome Sequence Archive of the Beijing Institute of Genomics (BIG) Data Center, BIG, Chinese Academy of Sciences, under accession code HRA006107 and are publicly accessible at http://bigd.big.ac.cn/gsa-human.
- All original code has been deposited at Mendeley Data and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this work paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Patients and samples

This study was approved by the ethics committee of Peking University People's Hospital (No. 2020PHB292-01). Patents and related samples were selected from specimen repository in our center between 2013 and 2018 according to follow criteria: 1. Clinical stage I lung cancer; 2. Received radical surgery; 3. Recurrence or death in 3 years. We excluded subjects with a prior history of cancer and antibiotic use (less than 1 month) or neoadjuvant therapy before surgery. Twenty-nine patients were included as recurrence group (R) and after matching for various clinicopathologic variables, 29 patients without known recurrence (NR) was chosen to the recurrence group (16S cohort). Detailed clinical and pathologic information on the patients is presented in Table 1. Tumor and normal lung specimens were frozen in a liquid nitrogen tank immediately after resection in a sterile environment, and then transferred to −80°C until processing for DNA extraction. Another cohort of 63 treatment-naive lung cancer patients with R and NR were enrolled in this study according to above-mentioned criteria (cmDNA cohort). Detailed clinical and pathologic information on the patients is presented in Table S5. Plasma samples were collected before surgery and transferred to −80°C until processing for DNA extraction. The samples used in Figure S1I were obtained from biobank of Jiangsu Cancer Hospital (Jiangsu Institute of Cancer Research & The Affiliated Cancer Hospital of Nanjing Medical University).

Cell lines and reagents

All cell lines (A549, NCI-H1299, LLC and RAW 264.7) were obtained from the Cell Bank of the Chinese Academy of Sciences (Shanghai, China). LLC and RAW 264.7 cells were cultured in Dulbecco's modified Eagle's medium (DMEM, Gibco, 2125039), and others were cultured in RPMI 1640 medium (Gibco, 2122752). All media contained 10% fetal bovine serum (FBS) and 100U per mL of penicillin–streptomycin. All cell lines were grown at 37°C with 5% CO2. Sodium butyrate (Solarbio, IS0190) was used for butyrate-related experiments. *Roseburia intestinalis* (DSMZ-14610) was purchased from Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (Braunschweig, Germany) and grown anaerobically at 37°C in modified YCFA medium.[45] The number of live bacteria (colony-forming units, CFU) was counted according to the absorbance at 600 nm (A600).

Bone marrow–derived macrophage (BMDM) isolation and differentiation

BMDMs were isolated as previously described with minor modification.[46] Six-week-old C57BL/6 mice were euthanized and sterilized with 75% ethanol. The femurs were dissected and muscle tissue was removed using scissors. Then, the bones were cut from both ends and flushed with medium using a 1-mL syringe to extrude the bone marrow into a culture dish. Bone marrow cells were cultured in DMEM containing 10% FBS and 20 ng/mL macrophage colony stimulating factor (M-CSF) for 6 days to obtain BMDMs. Then, IL4 (20 ng/mL) was added to generate full alternatively activated M2-BMDMs.

Lung cancer derived organoids culture

Human lung cancer samples were transported to the laboratory on ice within 0.5 h of removal from the patients in cold Hank's balanced salt solution (HBSS) with antibiotics (Lonza, Basel, Switzerland). Samples were washed three times with cold HBSS with antibiotics, then minced and shaken in digestion solution containing 0.001% DNase (Sigma-Aldrich, MO, USA), 1 mg/mL collagenase/dispase (Roche, IN, USA), 200 U/mL penicillin, 200 mg/mL streptomycin in DMEM/F12 medium at 37°C for 0.5 h with intermittent agitation. After incubation, the suspensions were repeatedly triturated by pipetting and passed through 100 μm cell strainers (BD Falcon, CA, USA). The strained cells were centrifuged at 1000 rpm for 5 min, and the pellet were washed with 1 mL cold HBSS. Then centrifugation at 1000 rpm for 1 min, remove supernatant, the organoids were embedded into matrigel (356231, Corning) at a density of 1 organoid per μL and plated in a 24 Well Cell Culture Plates (3337, Corning) for 15 min. After gelation, 600 μL MBM serum-free medium (DMEM/F12, Gibico) supplemented with 50 ng/mL human EGF (AF-100-15, Peprotech), 100 ng/mL FGF10 (100-26, Peprotech), 100 ng/mL FGF4 (100-31, Peprotech), B27(17504-044, Gibco), and 1% penicillin/streptomycin (Gibco, OK, USA) was added to the well. The medium was changed every 3 days, and the organoids were passaged after 10–14 days.

Animals

All procedures were approved and performed in accordance with relevant guidelines of the Animal Care Committee of Peking University People's Hospital (No. 2018PHC066). Male C57BL/6 mice (4–6 weeks, 18-20g) were purchased from Charles River (Beijing, China) and were maintained in the barrier facility.

For LLC mice metastasis model, mice were aerosolized with vancomycin (10 mg/ml) and neomycin (20 mg/ml) for 1week, and randomly divided into three groups (4 mice in each). Then, mice were aerosolized with saline or butyrate (40mM or 200mM) every 2 days for 3 weeks. All mice were intravenous tail vein injected with 1×10^6 LLC cells 1 day after the first treatment, and the number of melanotic foci was counted 3 weeks later.

Another group of mice were aerosolized with vancomycin (10 mg/ml) and neomycin (20 mg/ml) during the first week and then randomly divided into four groups (8 mice in each). Starting from the sixth day of aerosolized antibiotics, mice received intraperitoneal injections of either clodronate liposomes or control liposomes every 5 days. Subsequently, mice were aerosolized with saline or butyrate (40mM) every 2 days for a duration of 3 weeks. On the first day after the initial treatment, all mice were intravenously injected with $1x10^6$ LLC cells via the tail vein. The number of metastatic tumor foci was then counted after 4 weeks.

An orthotopic lung cancer model was generated in C57BL/6J mice as described previously with small modification.[47] LLC cells was pretreated with DMEM with or without butyrate (1mM) for 24h. Then, 5×10^5 LLC cells mixed with Matrigel (BD, 356324) were injected into the left lungs of mice. All mice were euthanized at d 21 after the injection of tumor cells.

For subcutaneous xenograft mouse model, 1×10^6 LLC cells were subcutaneously injected into one flank of each mouse. Seven days after subcutaneous inoculation, the mice were randomly divided into different groups. *R. intestinalis* resuspended in PBS was adjusted to 10^6 CFU/100ul. *R. intestinalis* suspension or PBS was given by multipoint intratumoral injection every three days. The mice were euthanized before the tumors were dissected 2 weeks later.

METHOD DETAILS

DNA extraction and 16s gene sequencing

16S rRNA sequencing was performed by the Microbial Genome Research Center (IMCAS, Beijing, China). Briefly, DNA of the tissue sample was extracted and the V3 and V4 regions of the bacterial 16S rDNA gene were amplified. Purified amplicons were pooled in equimolar amounts and paired-end sequenced (2 × 250) on an Illumina MiSeq platform according to standard protocols. FLASH software (version 1.2.11, https://ccb.jhu.edu/software/FLASH/index.shtml) was used to merge paired-end reads from next-generation sequencing.[48] Low-quality reads were filtered by FASTX Toolkit (version 1.2.11, http://hannonlab.cshl.edu/fastx_toolkit/), and chimera reads were removed by USEARCH (version 11) program's UCHIME command and the "GOLD" database. After a random selection of 20,000 reads, the taxonomical classification of reads was determined using the RDP classifier (version 2.7) to generate the composition matrices at the level of the phylum to the genus.[49] A bootstrap value >0.8 was considered as high-confidence taxonomy assignment, while low-confidence sequences were labeled as unclassified assignment. Alpha diversity in our samples were calculated and displayed by vegan R package. Principal coordinate analysis (PCoA) was performed to visualize the Beta diversity between different groups. The linear discriminant analysis (LDA) effect size (LEfSe) method was used to detect microbial biomarkers (|LDA| score >2.5 and $P < 0.05$) among different groups.[50] Lasso regression model was used to further selection of microbial biomarkers. The genus predicting score was generated as follows: $genus\ predicting\ score = \beta_1 x_1 + \beta_2 x_2 + + \beta_i x_i$ where β_i is the coefficient of each genus and x_i is the relative abundance of each genus.

Circulating microbiome DNA sequencing and analysis

Whole blood was collected in EDTA tubes after skin surfaces were sterilized twice and processed immediately. Plasma and cellular components were separated by centrifugation at 1600*g* for 10 min at 4°C. Plasma was centrifuged a second time at 16,000 g at 4°C to remove any remaining cellular debris and stored at −80°C until the time of DNA extraction. NGS cfDNA libraries were prepared for whole genome sequencing using 10 to 250 ng of cfDNA. Briefly, the Qubit dsDNA HS Assay Kit was used to measure cfDNA concentrations according to the manufacturer's recommendations. Then, genomic libraries were prepared using the VAHTS Universal DNA Library Prep Kit for Illumina V3. Whole genome libraries were sequenced using 100-bp paired-end runs on the DNBSEQ-T7, which was performed by Geneplus-Beijing Institute (Beijing, China). All sequence reads were first mapped to reference sequence hg19 (Human Genome version 19) using Bowtie2 (v2.3.5.1) with default parameters. Reads that mapped to human genome were removed using Samtools software. The filtered reads were mapped to NCBI microbial reference genome databases using k-*mer*-based algorithm with Kraken. Relative abundance at bacterial genus level were estimated by Braken with recommended parameters. We used the MaAslin2 software[40] to get genera with top predictive ability in discovery set with q value < 0.25. Random forest model with selected genera as input was constructed with the caret package and the randomForest R package.[41] The receiver operating characteristics (ROC) curve and class predictions were generated by pROC R package.[42]

RNA sequencing and analysis

High-throughput sequencing was performed by Annoroad Gene Technology Corporation (Beijing). Sequencing libraries were generated using NEBNext Ultra RNA Library Prep Kit for Illumina (#E7530L, NEB, USA) following the manufacturer's recommendations and index codes were added to attribute sequences to each sample. and then analyzed for quality on Agilent Bioanalyzer 2100 system

(Agilent Technologies, CA, USA). After cluster generation, the libraries were sequenced on an Illumina platform and 150 bp paired-end reads were generated. Reads were aligned to the human genome (Ensembl release 82- GRCh38) using HISAT2 v2.1.0 and default parameters. Reads Count was counted by HTSeq v0.6.0,[43] and FPKM was then calculated to estimate the expression level of genes. DESeq2[44] was used for differential gene expression analysis. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of differential expression genes were performed, with adjusted $p < 0.05$ and absolute fold change (FC) > 2. The differentially upregulated genes were used as a gene signatures in GSEA analyses. GESA was performed using GSEABase R package.

In addition, the expressions of H19 and MMP15 and their relationship with disease-free survival in TCGA-NSCLC cohort were analyzed through Gene Expression Profiling Interactive Analysis 2 (http://gepia2.cancer-pku.cn/).

RNA extraction and qRT-PCR analysis

RNA extraction and qRT-PCR were performed as described previously.[51] The gene expression levels were normalized to internal controls (GADPH or 18S), and the $2^{-\Delta\Delta Ct}$ method was used to calculate the fold changes in expression. The primers used are indicated in Table S9.

Construction of small interfering RNAs (siRNAs) and transfection of cells

SiRNAs were designed and synthesized by Tsingke Biological Technology (Beijing, China). Lung cancer cells were seeded in 6-well plates and 24 h later, at a confluence of 60–70%, were transfected with specific siRNA (100 nM) or control siRNA (100 nM) using Lipofectamine RNAiMAX according to the manufacturer's protocol (Invitrogen, 13778100). All siRNA sequences used are listed in Table S10.

Cell proliferation, migration, and invasion

Cells were seeded into a 96-well plate and incubated in culture medium with variable concentrations of butyrate for 24 h. Subsequently, the cells were labeled using a Cell Counting Kit-8 (CCK8) (Biosharp, BS350A) for 2 h. The absorbance of each well was measured with a microplate reader set at 450nm. Colony forming assays were performed to evaluate the clonogenic ability of the cells. Wound healing assays and transwell assays were performed to evaluate the migration and invasion capabilities, respectively. A549 or H1299 cells were grown in 24-well plastic dishes and treated with butyrate or control for 24h. Then, 4×10^4 cells in serum-free medium were seeded in the transwell migration chambers (Millipore, PIEP12R48) with or without diluted Matrigel (BD, 356324). Medium containing 10% FBS was added to the lower chambers. For LLC, cells were resuspended in 200μL serum-free DMEM and seeded to the upper chamber. CM (500μL) from BMDMs or RAW264.7 cultured in the presence or absence of butyrate was added to the lower chamber. Migrated and invaded cells were stained with crystal violet and were then counted using a light microscope. Each experiment was performed in triplicate.

Chromatin immunoprecipitation (ChIP) and qPCR

The ChIP assay kit (Beyotime Institute of Biotechnology, P2078) was used according with the manufacturer's protocol. In briefly, A549 cells were cross-linked with 1% formaldehyde for 10 min at 37 °C, collected in SDS lysis buffer. After sonication, the cell extract was incubated with anti-HADC2 (12922-3-AP), anti-Histone H3 (acetyl K27, ab177178) or control anti-IgG antibody overnight at 4°C. Protein A + G-agarose beads were added and incubated for 1 h at 4 °C. After reversing the cross-links and purification, DNA was used for qPCR reactions. The ChIP PCR primers used were indicated in Table S9.

Western blot analysis

Cells were lysed RIPA buffer (Cell Signaling Technology, 9806S) containing a protease inhibitor and phosphatase inhibitor cocktail (Sigma, MSSAFE-1VL). The concentration of the extracted protein was measured using the BCA method, and samples were subjected to SDS-PAGE. Proteins were transferred to polyvinylidene fluoride membranes (Millipore, SAMP2GVNB), which were blocked at room temperature for 1h in 5.0% non-fat milk and incubated overnight at 4°C with primary antibodies. Membranes were washed three times with TBST and incubated with the corresponding HRP-conjugated secondary antibodies for 2 h at room temperature. Protein detection was performed using a chemiluminescence system (Bio-Rad, USA).

Histology immunohistochemistry (IHC) and Immunofluorescence (IF)

Tissue samples were fixed in PBS containing 10% neutral-buffered formalin. Paraffin-embedded sections (5μm) were stained with hematoxylin and eosin. IHC was applied to detect protein expression according to a standard protocol.[52] Tissue sections were deparaffinized and rehydrated prior to antigen retrieval in citrate buffer. The sections were then stained with antibodies against *Arg1*(1:2000; Abcam, ab233548), CD206 (1:2000, Proteintech, 18704-1-AP), Arg1 (1:500, Servicebio, GB11285), MMP15 (1:300, Immunoway, YT2911) or F4/80 (1:500, CST, 70076S). For IF, monoclonal antibodies were incubated with the corresponding fluorescent secondary antibodies (Servicebio, GB21303, GB25301, 1:300). Positive and negative control sections were always included in the IHC staining and IF protocol.

Conditioned medium (CM) preparation

Macrophage polarization was obtained by culturing cells in DMEM medium supplemented with 10% FBS and 20 ng/mL IL4 with or without butyrate for 48h. Polarized RAW264.7 or BMDM cells were then incubated in serum-free medium for 24h, after which culture

supernatants were collected as conditioned medium (CM). CM was centrifuged at 3,000 rpm to remove debris and then stored at −80°C.

Cytokine detection in supernatant of cell culture media

CM from RAW 264.7 treated with or without butyrate and IL4 was collected for cytokine detection using Quantibody Mouse TH17 Array 1 (RayBiotech, Guangzhou, China). Briefly, take out the glass slide from the box, and let it equilibrate to room temperature inside the sealed plastic bag for 30 min. Remove slide from the plastic bag, peel off the cover film, and let it air dry for another 1.5 h. Reconstitute the Cytokine Standard Mix (lyophilized) by adding 500 μL Sample Diluent to the tube. For best recovery, always quick-spin vial prior to opening. Then, dissolve the powder thoroughly by a gentle mix. Labeled the tube as Std1. Label 6 clean microcentrifuge tubes as Std2 to Std7. Add 200 μL Sample Diluent to each of the tubes. Pipette 100 μL Std1 into tube Std2 and mix gently. Perform 5 more serial dilutions by adding 100 μL Std2 to tube Std3 and so on. Add 100 μL Sample Diluent to another tube labeled as CNTRL. Do not add standard cytokines or samples to the CNTRL tube, which will be used as negative control. For best results, include a set of standards in each slide. Add 100 μL Sample Diluent into each well and incubate at room temperature for 30 min to block slides. Decant buffer from each well. Add 100 μL standard cytokines or samples to each well. Incubate arrays at room temperature for 1 h. Decant the samples from each well, and wash 5 times (5 min each) with 150 μL of 1X Wash Buffer I at room temperature with gentle rocking. Then, completely remove wash buffer in each wash step. Dilute 20x Wash Buffer I with H_2O. Decant the samples from each well, and wash 5 times (5 min each) with 150 μL of 1X Wash Buffer I at room temperature with gentle rocking. Completely remove wash buffer in each wash step. Dilute 20x Wash Buffer I with H_2O. Decant the samples from each well, and wash 5 times (5 min each) with 150μL of 1X Wash Buffer I and then 2 times with 150 μL of 1x Wash Buffer II at room temperature with gentle rocking. Completely remove wash buffer in each wash step. After briefly spinning down, add 1.4 mL of Sample Diluent to Cy3 equivalent dye-conjugated streptavidin tube. Mix gently. Add 80 μL of Cy3 equivalent dye-conjugated streptavidin to each well. Cover the device with aluminum foil to avoid exposure to light or incubate in dark room. Incubate at room temperature for 1 h. Decant the samples from each well, and wash 5 times (5 min each) with 150 μL of 1X Wash Buffer I at room temperature with gentle rocking. Completely remove wash buffer in each wash step. Disassemble the device by pushing clips outward from the slide side. Carefully remove the slide from the gasket. Place the slide in the Slide Washer/Dryer (a 4-slide holder/centrifuge tube), add enough 1x Wash Buffer I (about 30 mL) to cover the whole slide, and then gently shake at room temperature for 15 min. Decant Wash Buffer I. Wash with 1x Wash Buffer II (about 30 mL) and gently shake at room temperature for 5 min. Remove water droplets completely by gently applying suction with a pipette to remove water droplets. Do not touch the array, only the sides. Imaging: The signals can be visualized through use of a laser scanner equipped with a Cy3 wavelength (green channel) such as Axon GenePix or Innopsys Innoscan. Make sure that the signal from the well containing the highest standard concentration (Std1) receives the highest possible reading, yet remains unsaturated. Finally, data extraction can be done using the GAL file that is specific for this array along with the microarray analysis software.

Flow cytometry analysis

Flow cytometry was employed to analyze the peripheral blood macrophage population in mice. A total of 200 μL of whole blood was collected from the eye using EDTA blood collection tubes. Subsequently, 1 mL of red blood cell lysis buffer was added to ensure thorough mixing of the lysis buffer with the whole blood. The mixture was kept in the dark at room temperature for 10 min. Following this, centrifugation (500g, 5min, 4°C) was performed, and the supernatant was carefully discarded. Another centrifugation step was carried out with 1 mL of PBS at 500g for 5 min at 4°C, and the remaining lysis buffer and red blood cells were removed. Subsequently, fluorescently conjugated antibodies, CD11b (BD Biosciences, 561691) and F4/80 (BD Biosciences, 563899), were added to individual tubes and sample tubes, followed by incubation in the dark at room temperature for 15 min. Finally, the samples were analyzed using a flow cytometer.

QUANTIFICATION AND STATISTICAL ANALYSIS

The associations between variables were statistically analyzed using χ^2 test for categorical variables and unpaired sample Student's *t* test or Wilcoxon test for continuous variables. The Kruskal-Wallis test was used for comparison of multi-groups. RFS was defined as the time from surgery to the date of disease recurrence or death. Analyses of RFS were performed using the Kaplan–Meier method and log rank test. Two-sided P-values <0.05 were considered statistically significant. All analyses were performed in R software (version 3.6.1; R Development Core Team), GraphPad Prism 8 (GraphPad Software, Inc.), and SPSS software (version22.0; IBM Corporation, Armonk, NY, USA).