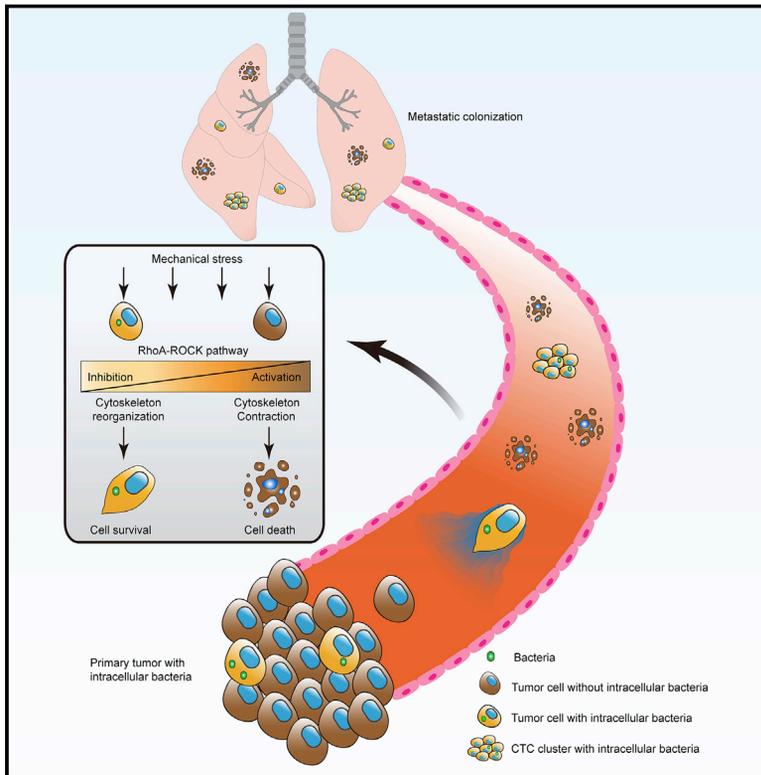


# Tumor-resident intracellular microbiota promotes metastatic colonization in breast cancer

## Graphical abstract



## Authors

Aikun Fu, Bingqing Yao, Tingting Dong, ..., Yajing Guo, Nan Li, Shang Cai

## Correspondence

caishang@westlake.edu.cn

## In brief

Tumor-resident intracellular bacteria enhance survival of circulating tumor cells by cytoskeleton reorganization. They promote metastasis but are not required for primary tumor growth in a murine breast cancer model.

## Highlights

- Conserved intracellular bacterial profile is detected in murine and human breast cancer
- Perturbation of intracellular bacteria reduces metastasis but not primary tumor growth
- Intracellular bacteria reorganize actin cytoskeleton in circulating tumor cells (CTCs)
- Intracellular bacteria promote CTC survival by enhancing resistance to mechanical stress



## Article

# Tumor-resident intracellular microbiota promotes metastatic colonization in breast cancer

Aikun Fu,<sup>1,2,3,6,10</sup> Bingqing Yao,<sup>7,1,2,3,10</sup> Tingting Dong,<sup>7,1,2,3,10</sup> Yongyi Chen,<sup>8</sup> Jia Yao,<sup>9</sup> Yu Liu,<sup>9</sup> Hang Li,<sup>5</sup> Huiru Bai,<sup>7,1,2,3</sup> Xiaoqin Liu,<sup>1,2,3,7</sup> Yue Zhang,<sup>7,1,2,3</sup> Chunhui Wang,<sup>1,2,3,4</sup> Yajing Guo,<sup>1,2,3</sup> Nan Li,<sup>5</sup> and Shang Cai<sup>1,2,3,4,6,11,\*</sup>

<sup>1</sup>Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, Zhejiang, China

<sup>2</sup>School of Life Sciences, Westlake University, Hangzhou, Zhejiang, China

<sup>3</sup>Key Laboratory of Growth Regulation and Translational Research of Zhejiang Province, School of Life Sciences, Westlake University, Hangzhou, Zhejiang, China

<sup>4</sup>Westlake Disease Modeling lab, Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, Zhejiang, China

<sup>5</sup>Westlake University High-Performance Computing Center, Westlake University, Hangzhou, Zhejiang, China

<sup>6</sup>Institute of Biology, Westlake Institute for Advanced Study, Hangzhou, Zhejiang Province, China

<sup>7</sup>School of Life Sciences, Fudan University, Shanghai, China

<sup>8</sup>Department of Clinical Laboratory, Zhejiang Cancer Hospital, Hangzhou, Zhejiang Province, 310000, China

<sup>9</sup>Department of Breast Surgery, First Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou, Zhejiang Province, China

<sup>10</sup>These authors contributed equally

<sup>11</sup>Lead contact

\*Correspondence: [caishang@westlake.edu.cn](mailto:caishang@westlake.edu.cn)

<https://doi.org/10.1016/j.cell.2022.02.027>

## SUMMARY

Tumor-resident intracellular microbiota is an emerging tumor component that has been documented for a variety of cancer types with unclear biological functions. Here, we explored the functional significance of these intratumor bacteria, primarily using a murine spontaneous breast-tumor model MMTV-PyMT. We found that depletion of intratumor bacteria significantly reduced lung metastasis without affecting primary tumor growth. During metastatic colonization, intratumor bacteria carried by circulating tumor cells promoted host-cell survival by enhancing resistance to fluid shear stress by reorganizing actin cytoskeleton. We further showed that intratumor administration of selected bacteria strains isolated from tumor-resident microbiota promoted metastasis in two murine tumor models with significantly different levels of metastasis potential. Our findings suggest that tumor-resident microbiota, albeit at low biomass, play an important role in promoting cancer metastasis, intervention of which might therefore be worth exploring for advancing oncology care.

## INTRODUCTION

The host microbiota, as a nonnegligible body component, is a crucial mediator in modulating cancer susceptibility and tumor progression in addition to the well-known genetic, epigenetic, and stromal microenvironment elements. These microbial organisms exert their functions notably through indirect pathways (including metabolites and the immune system) on distant or proximal tumor tissues, particularly in colorectal cancers where they are in intimate contact with the gut microbiota (Garrett, 2015; Xavier et al., 2020). However, in recent years, there are emerging lines of evidence that microbes are also integral components of the tumor tissue itself in much broader cancer types beyond colorectal cancer, such as pancreatic cancer, lung cancer, breast cancer, and others, which were originally thought to be sterile (Flemer et al., 2017; Jin et al., 2019; Nejman et al., 2020; Pushalkar et al., 2018; Riquelme et al., 2019; Urbaniak et al., 2016). Clinically, cohort studies have suggested that features of the tissue-resident microbiota correlate with cancer risks (Xuan et al., 2014), pathological types (Banerjee et al., 2018; Banerjee et al., 2015; Buchta Rosean et al., 2019; Hieken et al., 2016; Nejman et al., 2020), can-

cer prognosis (Riquelme et al., 2019), and treatment responses (Geller et al., 2017; Nejman et al., 2020; Yu et al., 2017). Nonetheless, tissue-resident microbiota samples are often of very low biomass with serious host contamination and potential environmental noise, which frequently obscure the findings (Davis et al., 2018; de Goffau et al., 2018; Jervis-Bardy et al., 2015; Kim et al., 2017; Laurence et al., 2014; Salter et al., 2014). Studies using patient-derived xenograft (PDX) models suggest that the intratumor microbiota can persist as the tumor progression in immunodeficient mice (Bullman et al., 2017). Administration of exogenous bacteria through tail vein impairs tumor chemosensitivity (Geller and Straussman, 2017; Yu et al., 2017) and promotes tumor progression (Parhi et al., 2020).

Whereas a growing body of evidence indicates the importance of intratumor microbiota, whether the low-biomass tumor-resident microbiota in its physiological homeostatic state plays any significant biological roles in spontaneous tumor progression is still an open question. An animal spontaneous tumor model that can recapitulate the human tumor microbiota will largely impel the functional investigation of tumor-resident microbiota and clarify its significance in tumor progression.



In the present work, we report an extensive characterization of such a spontaneous murine breast tumor (BT) model mouse mammary tumor virus-polyoma middle tumor-antigen (MMTV-PyMT) with significant amounts of intracellular bacteria, resembling that in human breast cancer. Under the physiological condition, these intracellular bacteria can travel through the circulation system along with the cancer cells and play critical roles in metastatic colonization. Mechanistically, the intracellular bacteria modulated the host-cell actin network and promoted cell survival against fluid shear stress in the circulation. Our study revealed that the intracellular bacteria can directly regulate the host-cell viability during tumor progression, which potentially could have strong implications for future cancer treatment.

## RESULTS

### Spontaneous murine BT contains a significant amount of live bacteria

As the presence of intratumor microbiota was reported in human cancers, we seek for an animal tumor system suitable for interrogation of tumor-resident microbiota functions. An MMTV-PyMT spontaneous murine BT model was selected for the intratumor microbiota characterization. Due to the low biomass challenge of intratumor microbiota (de Goffau et al., 2018), we optimized the qPCR protocol at multiple key steps (described in STAR Methods) (Figures 1A and S1A–S1C) and achieved a detection sensitivity of  $5 \times 10^3$  equivalent bacteria per gram of tissue (Figure 1B). With this method, we detected a median  $1.7 \times 10^4$  equivalent bacteria per gram (N = 18) of tissue in the normal mouse breast, and a nearly ten-fold higher bacteria load in the tumor tissue ( $1.35 \times 10^5$  equivalent bacteria per gram, N = 102,  $p < 0.0001$ ), both significantly higher than non-template control (NTC) and environment background control (EBC) ( $p < 0.0001$ ). The bacteria density in tumor tissue remained relatively constant as the tumor size increased (Figure 1C). These intratumor bacteria were culturable with various methods (Figure S1D), showing a median of 100 colony-forming units (CFU)/gram for normal breast tissue and 297 CFU/gram (Figures 1D and 1E) for tumor tissue, proportional to the qPCR result (Figure S1E). The isolated bacteria were enriched in *Staphylococcus*, *Lactobacillus*, *Enterococcus*, and *Streptococcus*, consistent with human breast tissues (Nejman et al., 2020; Thompson et al., 2017; Urbaniak et al., 2014; Urbaniak et al., 2016), and they largely overlapped at the genus level between normal and PyMT tumor, indicating a similar origin.

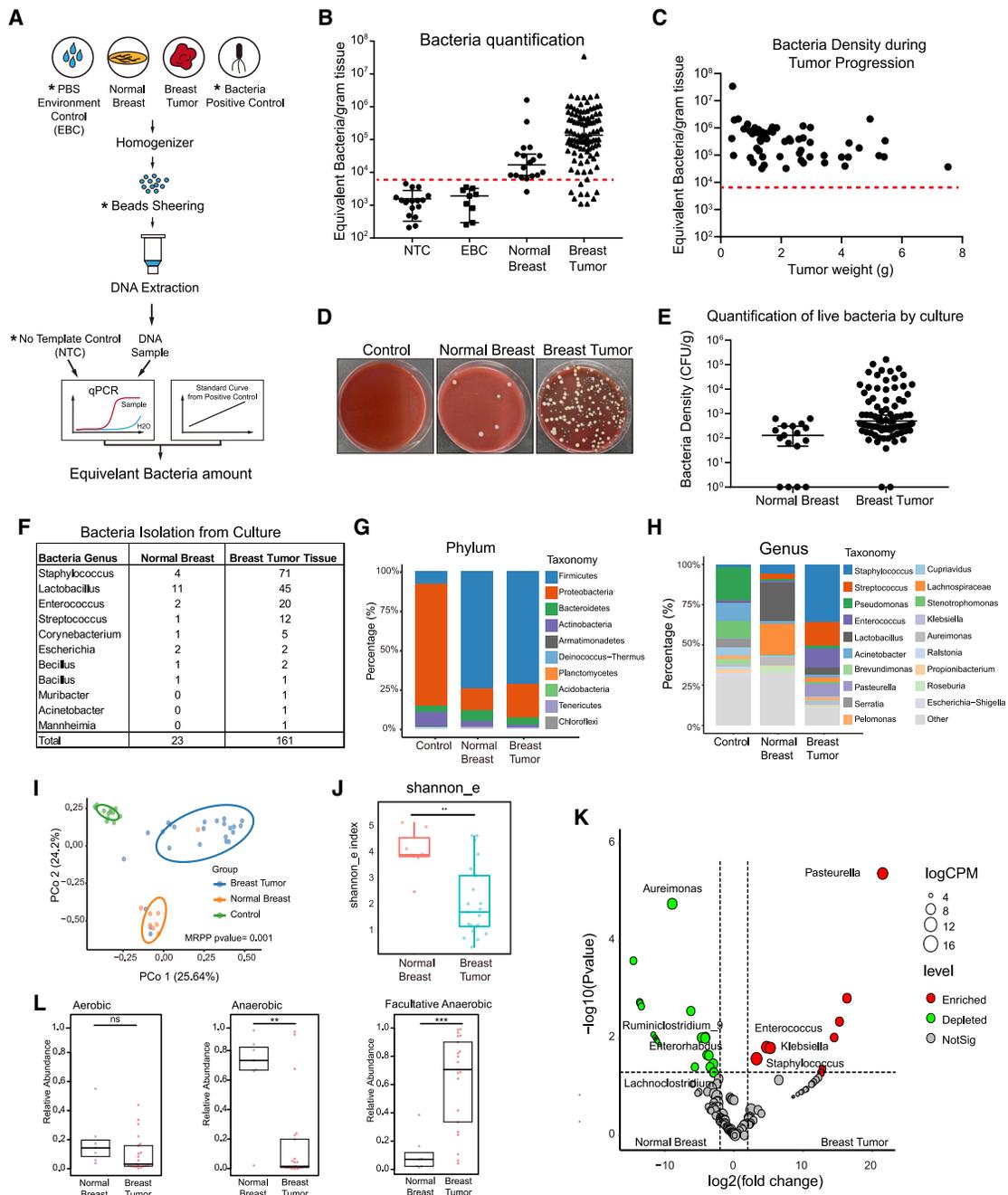
We further attempted to comprehensively profile the composition of the tissue-resident microbiota by 16S sequencing. Due to the challenges of the low biomass (Davis et al., 2018; de Goffau et al., 2018; Jarvis-Bardy et al., 2015; Kim et al., 2017; Laurence et al., 2014; Salter et al., 2014), we optimized the 16S library construction procedure by adding a biotin enrichment step to reduce the nonspecific genome sequences, and we achieved a faithful detection of  $10^4$  equivalent bacteria/gram of tissue, a significantly higher sensitivity than conventional methods for gut microbiota (Figures S1F–S1H; also see STAR Methods). With that, we analyzed the microbiota in the normal and tumor tissue (Figures 1G–1H). The unsupervised clustering of microbiome analysis showed that the controls (NTC and EBC), normal

breast tissue, and BT tissue had distinct microbial communities (Figure 1I). The majority of microbes in the negative control samples were *Proteobacteria* (Figures 1G and S1I), while the tissue samples were enriched for *Firmicutes* (Figure 1G). Upon contamination exclusion, we found that BT tissue has a reduced alpha diversity compared with normal breast tissue (Figure 1J,  $p < 0.01$ ), indicating a selection and expansion of certain microbes. The major genera detected in BTs are highly consistent with our culture isolates; namely, *Staphylococcus*, *Enterococcus*, *Streptococcus*, and *Lactobacillus* form the major components (Figures 1H, S1J, and S1K), with distinct features in normal and tumor tissue (Figure 1K; Table S1). Interestingly, BT tissue showed a drastic decrease of anaerobes and a remarkable increase of facultative anaerobes (Figure 1L), indicating a dynamic oxygen microenvironment in the tumor. These data collectively demonstrate that PyMT spontaneous tumor contains significant amounts of live bacteria.

### Significant amounts of tumor-resident microbes reside in the cytosol

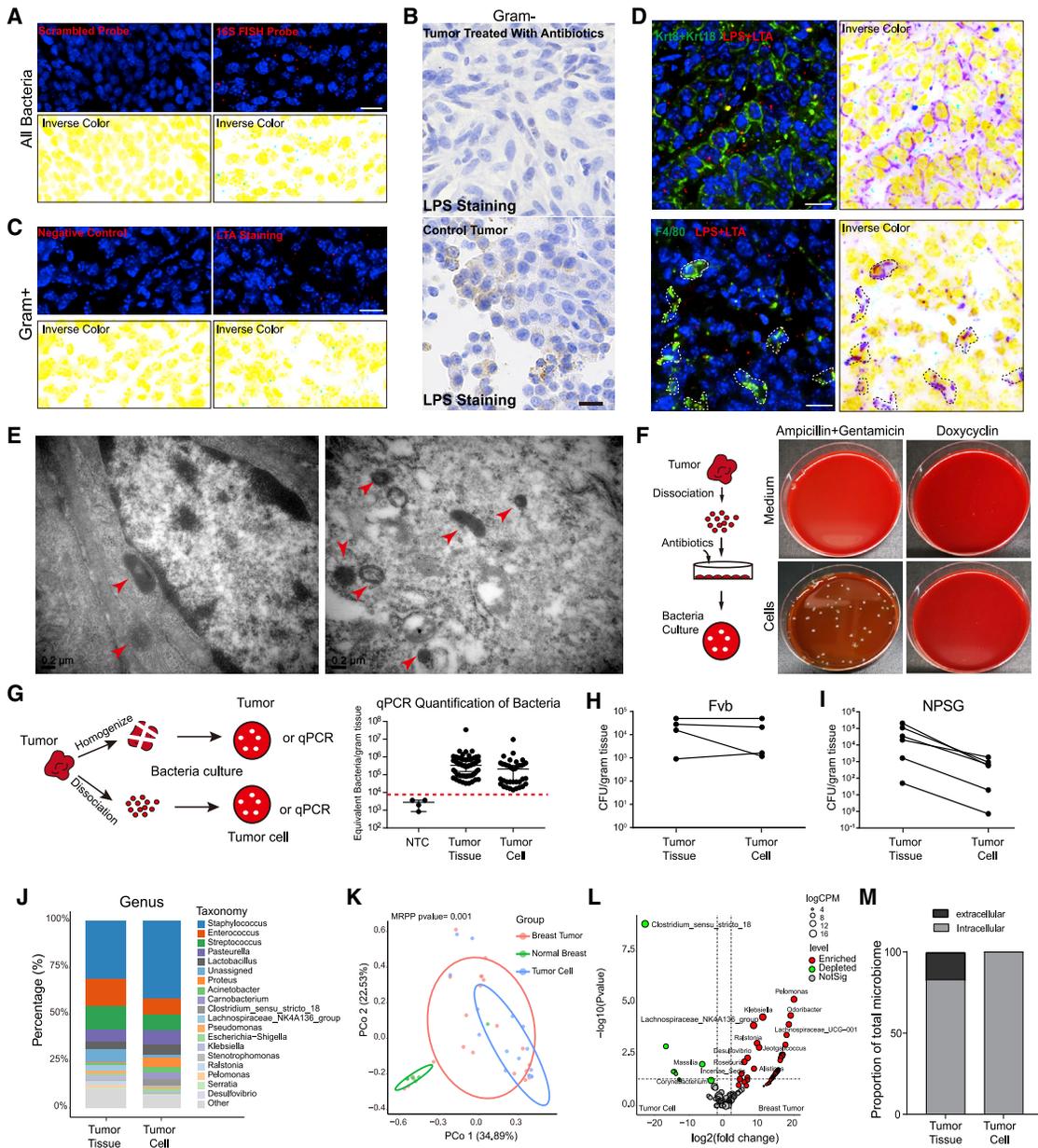
A recent study showed that intratumor bacteria in human cancers are predominantly present in the cytoplasm of both immune cells and tumor cells (Nejman et al., 2020). We further characterized the location of the microbes in the spontaneous PyMT BT. The 16S fluorescence *in situ* hybridization (FISH) analysis, lipopolysaccharides (LPS) staining (for gram-negative bacteria) and lipoteichoic acid (LTA) staining (for gram-positive bacteria) all showed the presence of bacteria at the perinuclear region as punctate dots (Figures 2A–2C). The density of bacteria in the tumor identified by RNA probe or antibody staining is much higher than qPCR quantification, which may result from inefficient bacteria DNA extraction in qPCR experiments. The identified bacteria preferentially inhabited the BT epithelia and less so stromal cells (Figures 2D, S2A, and S2B). High-resolution electron microscopy (EM) analysis showed that the majority of bacteria-like structures were identified in the cytosol rather than the extracellular space (estimated as 97.25% in cytosol, N = 218), and the bacteria-cell ratio was estimated to be 3% (218 bacteria out of 7201 scanned cells) (Figures 2E and S2C). These intracellular microbes were alive, and they survived upon cell impermeable antibiotics treatment (Ampicillin and Gentamicin) (Kumar et al., 2017) but not upon cell-penetrating doxycycline treatment (Figure 2F). This suggests that the bacteria we observed in tumor cell cytoplasm by staining and in EM were viable organisms.

We next tried to quantify the relative amount of extracellular and intracellular bacteria in the tumor by culturing on Columbia Blood Agar Base (CBA) plate the homogenized tumor tissue and the dissociated tumor cells. As enzyme-dissociated tumor cells underwent extensive washes, we assumed that the majority of the extracellular bacteria would be eliminated in the tumor cell sample. We found that the quantity of bacteria in tumor cells was not statistically different from the total bacteria in the same amount of tumor tissues (Figure 2G). This suggests that the number of extracellular bacteria in the PyMT tumor is minimal. Interestingly, we noticed that in the immunodeficient NPSG mice, there tend to be more extracellular bacteria components compared to immunocompetent Fvb mice (Figures 2H and 2I), indicating an immune involvement. To characterize the relative



**Figure 1. Microbiota is an integral component of MMTV-PyMT spontaneous breast tumor**

(A) Schematic diagram showing the workflow for qPCR quantification of tissue resident microbiota.  
 (B) qPCR quantification of the microbiota in normal breasts and PyMT spontaneous breast tumors. Data are presented as Median $\pm$ 95% CI.  $N_{NTC}=17$ ,  $N_{EBC}=9$ ,  $N_{normal}=18$ ,  $N_{tumor}=102$ . Red dashed line shows detection limit.  
 (C) Plot of bacteria density versus tumor weight by qPCR.  $N = 87$ .  
 (D) Representative pictures showing bacteria culture analysis of control (EBC), normal and PyMT tumor on CBA plate.  
 (E) Quantification of culturable bacteria of normal breast and PyMT tumor. Data are presented as Median  $\pm$  95% CI.  
 (F) Table showing the culture isolated bacteria genera in normal breast and PyMT spontaneous tumor.  
 (G and H) Stacked bar plot of the phylum and genus level relative abundance of bacteria communities in indicated samples.  
 (I) Unsupervised PCA plot via unifracs distance in indicated samples. Multiple Response Permutation Procedure (MRPP) test p value is 0.001.  
 (J) Shannon index of PyMT tumor and normal breast tissue. The Wilcoxon test  $p < 0.01$ .  
 (K) Volcano plot showing the signature bacteria in normal breast and PyMT tumor. FDR cutoff is 0.25, p value cutoff is 0.05.  
 (L) Characteristics analysis of microbiota in normal breast and PyMT tumor. Mann Whitney test P value. \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .  
 See also [Figure S1](#) and [Table S1](#).



**Figure 2. The intratumor bacteria predominantly reside in the cytosol of PyMT tumor cells**

(A–C) 16S FISH analysis (A), LPS IHC (B), and LTA staining (C) of MMTV-PyMT primary tumor section. Scrambled probe as the negative control. Red, 16S FISH probe or LTA staining. Blue, DAPI. Scale bar, 10µm.

(D) Co-staining of Keratin 8 + 18 (green)/LPS+LTA (red) (upper) and F4/80 (green)/LPS+LTA (red) (lower) to show the localization of Gram positive and negative bacteria in tumor. Scale bar, 10µm. Blue, DAPI.

(E) EM image of spontaneous MMTV-PyMT tumor showing bacteria structures within cytosol. Red arrows pointing to bacterial structures. Scale bar, 0.2µm.

(F) Bacteria culture of dissociated PyMT tumor cells treated with Ampicillin + Gentamicin or Doxycycline.

(G) Workflow and qPCR quantification of bacteria in tumor tissue and in tumor cells from PyMT spontaneous breast tumors. Data are presented as median ± 95% CI.

(H and I) the correlation between the bacteria in tumor tissue and in tumor cells in PyMT tumor on Fvb mice and NPSG mice.

(J) Stacked bar plot of relative abundance of the bacteria community in tumor tissues and in tumor cells.

(K) unsupervised PCA clustering analysis of the intratumor or intracellular bacteria using unifracs distance. MRPP test p value is 0.001.

(L) Volcano plot showing the differential bacteria strains in breast tumor tissues and in tumor cells. FDR cutoff is 0.25, p value cutoff is 0.05.

(M) Quantification of the relative abundance of extracellular bacteria and intracellular bacteria.

See also [Figure S2](#) and [Table S2](#).

proportion of extracellular bacteria, we performed 16S sequencing of tumor tissue and dissociated tumor cells. The microbiota profiles of tumor tissue and dissociated tumor cells at the genus level were largely the same (Figures 2J, 2K, S2D, and S2E). Most of the differential genera of bacteria are enriched in tumor tissue, such as *Klebsiella*, *Lachnospiraceae*, *Pelomonas*, and *Odoribacter* (Figure 2L; Table S2). We speculate that the bacteria genera enriched in tumor tissue are washed off during dissociation and most likely extracellular, which constituted ~20% of the relative abundance (Figure 2M).

### Elimination of tumor-resident microbes impedes metastasis but not primary tumor growth

Although studies have shown that gut microbiota contributes to the tumor progression in various tumor systems (Dapito et al., 2012; Li et al., 2016; Li et al., 2019; Sethi et al., 2018; Sivan et al., 2015; Zackular et al., 2013), the biological significance of intratumor microbiota remains largely unknown. There is a lack of proper tools to distinguish the function of gut and tumor-resident microbiota. We developed such a tool by testing different combinations of antibiotics and administration routes to selectively eliminate tumor-resident or gut microbiota (Figure 3A). We found that the administration of antibiotic cocktail (ATBx) (Iida et al., 2013; Pushalkar et al., 2018) through drinking water (DW) can efficiently eliminate both gut and tumor microbiota by  $10^4$ -fold and  $10^2$ -fold, respectively (Figures 3B and 3C). When ATBx is administered by intravenous (IV) injection through the tail vein, the tumor microbiota is eliminated while the gut microbiota is left intact (Figures 3B and 3C). Administration of the cell membrane penetrating antibiotic doxycycline through DW has a similar effect as ATBx-IV (Figures 3B and 3C). Most importantly, the ATBx-IV and Dox-DW treatment remained comparable to control group regarding the live bacteria quantity in the gut, the gut microbiota profile and the tumor immune profile (Figures S3A–S3D and S3F–S3I). With these different antibiotic treatment strategies, we addressed the distinct functions of gut microbiota and tumor microbiota in tumor progression. We found that ATBx-DW efficiently slowed down the tumor growth of spontaneous MMTV-PyMT BTs (total tumor weight  $15 \pm 0.5$ g in control versus  $11 \pm 0.8$ g in ATBx-DW,  $p < 0.001$ ), while eliminating tumor microbiota alone by ATBx-IV or Dox-DW did not affect total tumor weight at all (Figure 3D). This suggests that gut microbiota plays an important role in tumor growth, while the tumor microbiota does not.

At the advanced stage of MMTV-PyMT tumor development, most of the tumor bearing mice develop lung metastases (Guy et al., 1992). Interestingly, we noticed that the presence of microbiota in the primary tumor correlated weakly with the presence of lung metastases ( $r^2 = 0.3$ ,  $p < 0.05$ ), while the intracellular bacteria amount had a much strong linear correlation ( $r^2 = 0.6$ ,  $p < 0.01$ ) (Figure 3E). This suggests that the intracellular microbiota could be involved in tumor metastasis. We then analyzed the lung metastasis under different antibiotic treatment strategies and confirmed that eliminating tumor microbiota alone by ATBx-IV or Dox-DW reduced lung metastasis more than 3-fold ( $p < 0.05$ ) (Figure 3F). Since ATBx-DW eliminated both gut and tumor microbiota, it also showed an inhibition effect in tumor metastasis along with the total tumor weight reduction (Figure 3F). Notably, other macrolides antibiotics that selectively eliminate tumor microbiota

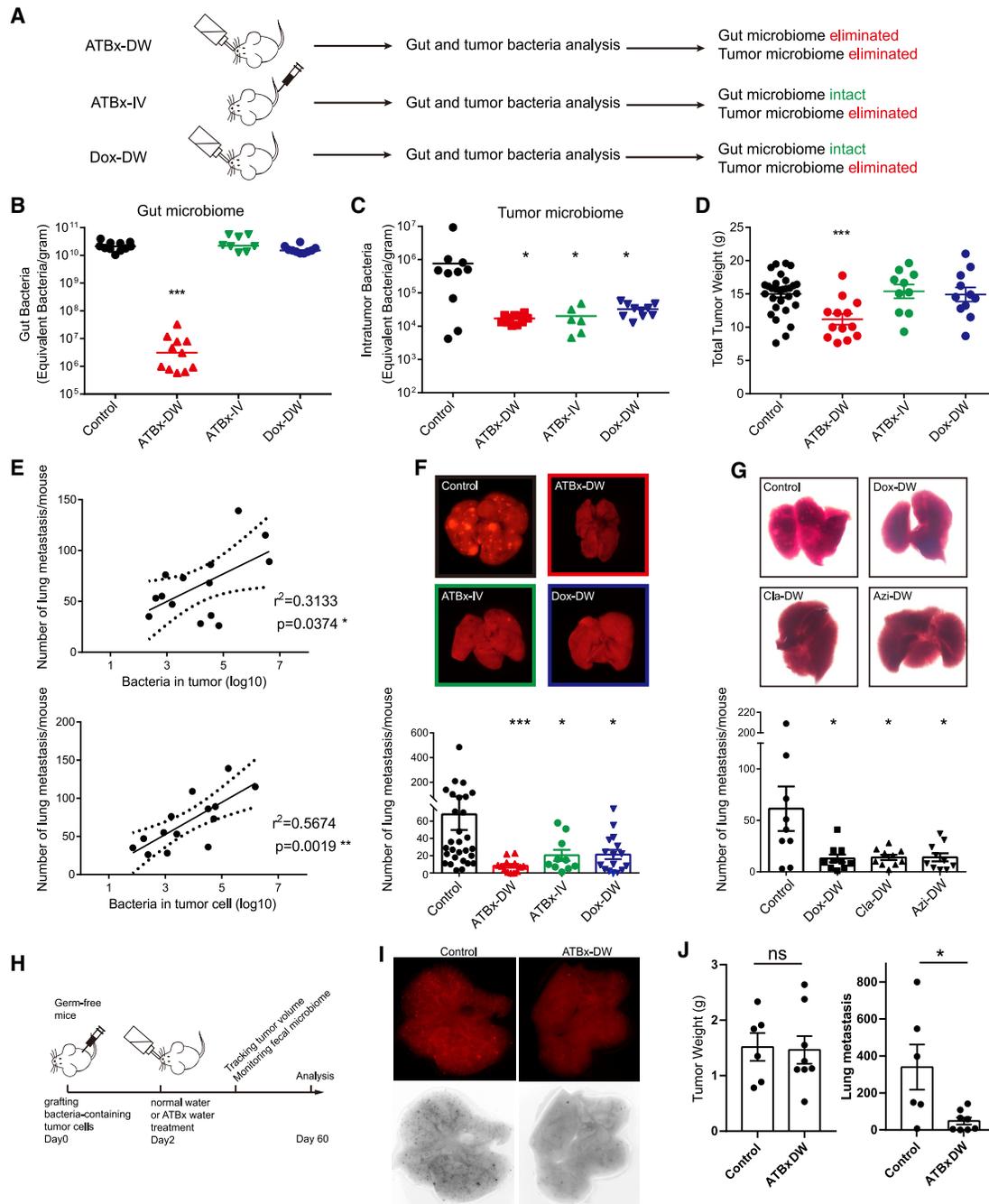
such as cell-membrane penetrating clarithromycin and azithromycin also led to a strong reduction in lung metastasis (Figure 3G) with no observable effect on primary tumor growth (Figure S3E). To completely exclude the influence of gut microbiota, we performed antibiotic treatment assay on germ free mice grafted with bacteria-containing PyMT tumors and observed consistent phenotypes (Figures 3H–3J, S3J, and S3L–S3N). Of note, the control tumor bearing germ-free mice showed drastic increase of metastatic foci compared with specific-pathogen-free (SPF) recipients, possibly due to their slow tumor growth and therefore extended experimental endpoint (Figure S3K), reinforcing a role of gut microbiota in tumor growth. These data collectively suggest that the intratumor bacteria are functionally essential for PyMT BT metastasis.

### Cancer cells carried over bacteria from primary tumor to metastatic site

As most of the bacteria we identified were intracellular, we speculated that during the metastatic process, the intracellular bacteria may be able to travel through the circulation system and migrate together with hosting tumor cells to the distal organ. This notion was supported by the 16S sequencing of the primary BT, metastasis adjacent lung containing early micro-metastasis (MAL) and lung macro-metastasis (LM), along with normal lung (NL) and normal breast (NB) tissue (Figures 4A–4C, S4A, and S4B), showing that the BT microbiota co-clustered with MAL sample very well, but was distinct from the NB and NL (Figures 4C, 4D, and S4C). The LM microbiota lies in between BT and normal tissues (Figures 4C and 4D). Further microbiota analysis showed that the aerobic bacteria components were increasing in the lung metastasis, while the facultative anaerobes were decreasing (Figure 4E), indicating an environment alteration in the lung such as the oxygen level.

We next tested the localization of bacteria in the lung metastases by 16S FISH analysis. Intriguingly, the tumor-adjacent normal lung tissue does not contain many bacteria, most of the identified bacteria were enriched in the metastatic foci (Figures 4F and S4D). This suggests that the metastatic tumor cells established a microenvironment that better supported the microbiota to thrive in. Further circulating tumor cell analysis in PyMT blood revealed that a small number of single circulating tumor cells were positive for bacteria (3 out of 206,  $n = 8$ ), but in contrast, a big portion of the circulating tumor cell clusters contained detectable bacteria (17 out of 34,  $n = 4$ ) (Figure 4G). There was an enrichment of bacteria in the circulating tumor clusters and in the lung mets compared with the tumor at primary site (Figure S4E), suggesting a beneficial role of bacteria specific for metastatic tumor cells. Recent years, emerging evidence has revealed that circulating tumor clusters have a better ability to survive and initiate distal organ metastasis (Aceto et al., 2014; Cheung et al., 2016). Given the common presence of bacteria in circulating tumor clusters, we propose that the intracellular bacteria can travel to the distal organ together with the host tumor cells offering the possibility to establish the metastatic microbiota themselves.

To visualize the dynamics of microbiota during the metastasis process, we plotted the abundance of various bacteria genera in different samples with the order NB, BT, MAL, LM as an indication of tumor progression process. Interestingly, unsupervised



**Figure 3. Tumor-resident microbiota are essential for PyMT spontaneous tumor metastasis**

(A) Schematic diagram showing various administration strategies of antibiotics and the influences on the gut and tumor microbiota. ATBx-DW, ATBx combined antibiotics via drinking water; ATBx-IV, ATBx combined antibiotics via iv injection; DOX-DW, Doxycycline via drinking water.

(B and C) qPCR Quantification of the bacteria abundance in the gut (B) and PyMT tumor (C) after various antibiotics treatment. Data are presented as median. Mann Whitney test \*\*\*  $p < 0.001$ , \*  $p < 0.05$ .

(D) Total tumor weight of spontaneous PyMT tumors after indicated antibiotics treatment. t test \*\*\*  $p < 0.001$ .

(E) Correlations between the number of lung metastasis and the intratumor or intracellular bacteria abundance quantified by culture.

(F) Quantification of lung metastasis number after indicated antibiotics treatment on MMTV-PyMT mTmG mice. Mann Whitney test \*  $p < 0.05$ . \*\*\*  $p < 0.001$ .

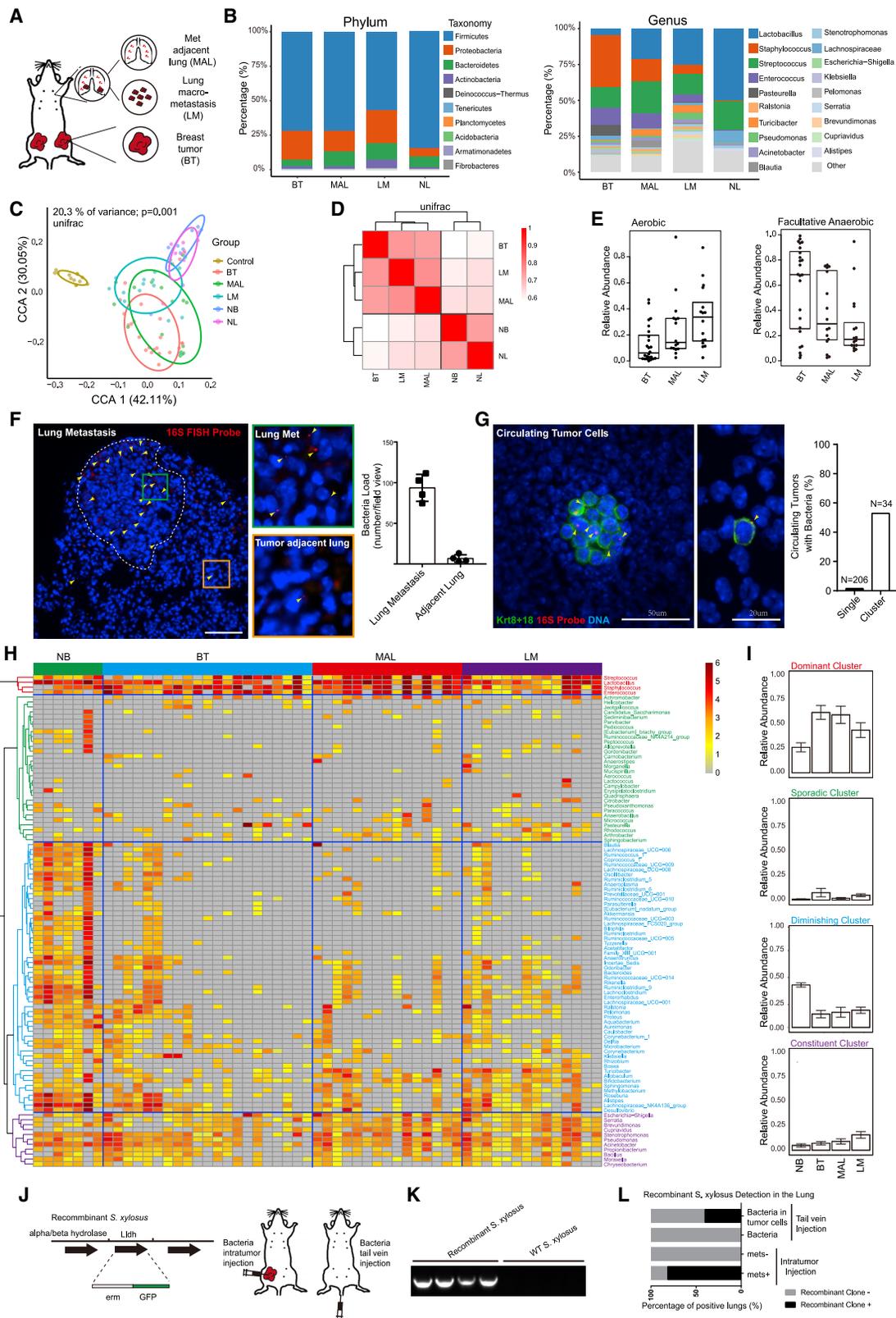
(G) Quantification of the lung metastasis number after clarithromycin and azithromycin treatment on MMTV-PyMT mice. Mann Whitney test \*  $p < 0.05$ .

(H) Schematic diagram showing antibiotic treatment of germ-free mice bearing PyMT-mTmG tumors.

(I) Fluorescence images of lung showing PyMT tumor cell metastases on germ-free mice in the presence or absence of antibiotic treatment.

(J) Tumor weight and lung metastases of primary PyMT tumors on germ-free mice with or without antibiotic treatment. \*  $p < 0.05$ . Data in this figure are all presented as mean  $\pm$  SEM unless otherwise indicated.

See also Figure S3.



(legend on next page)

clustering divided the bacteria into four categories with distinct dynamic patterns (Figures 4H and 4I; Table S3). The “dominant cluster” increased the abundance in the breast tissue, which persisted in micro-metastasis but decreased in macro-metastasis (Figures 4H and 4I; Table S3). We reason that this cluster of bacteria is likely to travel with metastasizing tumor cells and establishes the microbial environment early on, but as the metastatic growth processes it is influenced by the distal organ environment. The “sporadic cluster” contained some bacteria strains that only presented sporadically with no observable pattern. The “diminishing cluster” contains bacteria that have high abundance in the normal breast tissue but were gradually diminishing in the BT, micro-metastasis, and macro-metastasis. The “constituent cluster” includes low-abundance bacteria that were constantly detected in various samples (Figures 4H and 4I; Table S3). We suspect that the dominant cluster may be functionally involved in the tumor metastasis process and is therefore worthy of further research. Coincidentally, the bacteria strains within this cluster, *Streptococcus*, *Lactobacillus*, *Staphylococcus*, and *Enterococcus*, were all successfully isolated using our culture method.

To further test the hypothesis that bacteria can travel with the tumor cells in the circulation system, we engineered an erythromycin resistance element-GFP (*erm*-GFP) cassette into the *S. xylosus* genome to trace the bacteria spatially and temporally (Figure 4J). We found that when we introduced the recombinant *S. xylosus* directly into the tumor, we could successfully identify the recombinant clone in 80% of the lungs with metastases. In contrast, for those lungs without metastasis, no recombinant clones could be identified (Figures 4K, 4L, S4F, and S4G). In addition, if we introduced the recombinant *S. xylosus* through tail vein injection, we could only detect the recombinant clone in the lung when bacteria invaded tumor cells were injected, while bacteria injection alone, even at higher dose, fail to colonize in the lung (Figures 4K and 4L). These data suggest that bacteria can, but only together with tumor cells, travel through circulation system and colonize in distal organs.

### Certain intracellular bacteria promote metastatic colonization

Bacteria can invade the cytoplasm of eukaryotic cells and trigger cellular immunity along with other cellular behaviors. The effect

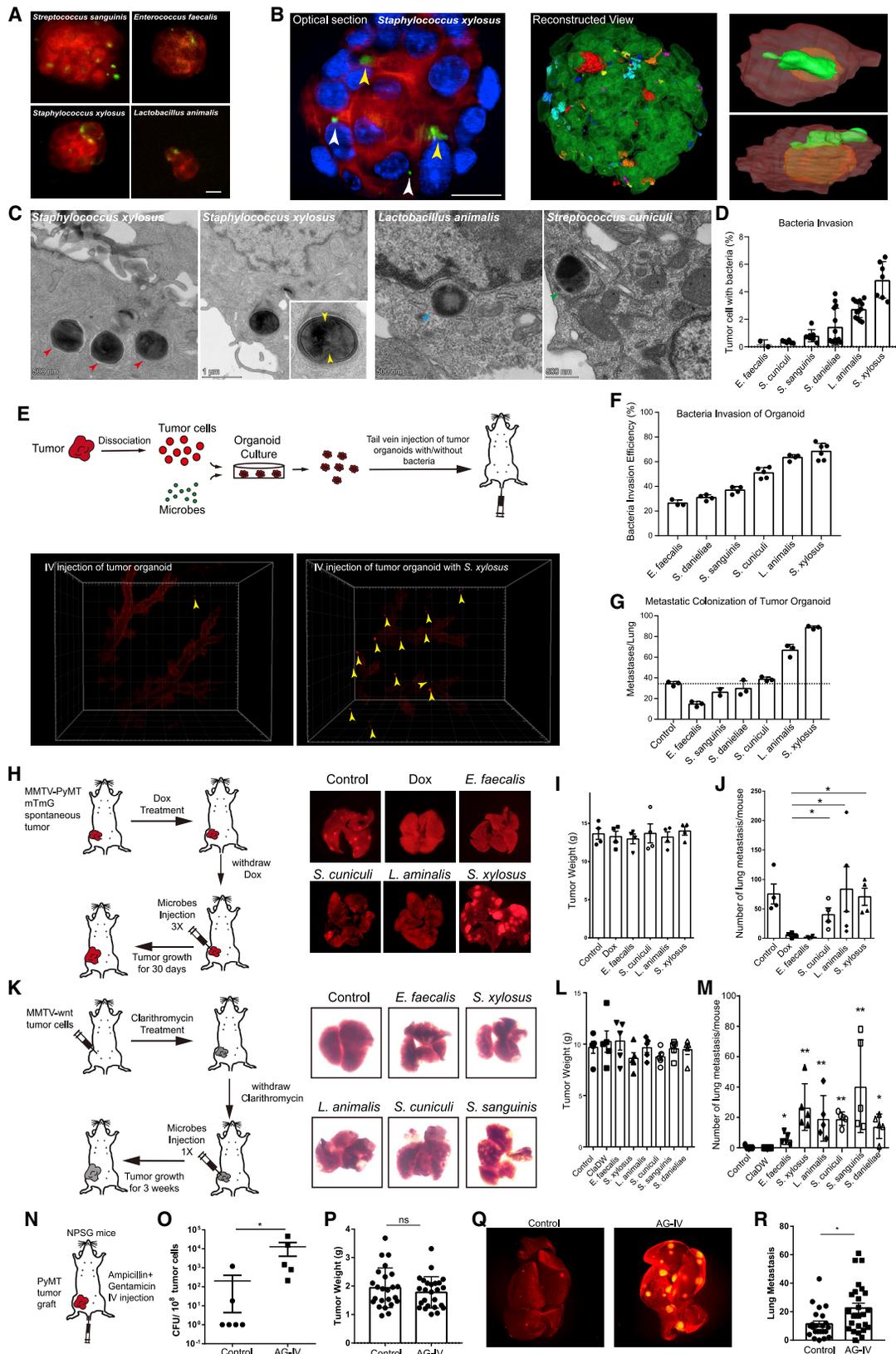
of bacteria on the target cells is highly dependent on the cell type and the bacteria strains (Strobel et al., 2016). To investigate how the tumor-resident bacteria influence the tumor cell activity, we established a Matrigel organoid (Cai et al., 2017) co-culture system in which individual bacteria strains labeled by green-fluorescent 5-(and 6)-carboxyfluorescein diacetate succinimidyl ester (CFSE) seeded with dissociated tumor cells (expressing tdTomato). Of note, we did not observe any influence of the isolated bacteria on the tumor cell colony growth (Figure S5A). However, fluorescence microscopy analysis showed that *Staphylococcus*, *Lactobacillus*, and *Streptococcus*, but not *Enterococcus*, can be efficiently incorporated into the organoid (Figure 5A) and exist as individual or clustered form in the cytosol (Figures 5B and S5B). Further EM analysis reinforced that those bacteria penetrated through the cell membrane and localize in the cytosol, but the vast majority of them were not in the endosome. Some bacteria even showed clear cleavage furrow, suggesting its active state in the cytoplasm (Figure 5C). Fluorescence-activated cell sorting (FACS) analysis of the invasion efficiency showed that *Enterococcus* barely invaded the tumor cells on the individual cell level in our in vitro culture system, while the other strains invaded the tumor cells more frequently (Figure 5D).

To ask whether the invasion of individual strain of bacteria could exert any function in the tumor cell colonization during metastasis, we injected tumor organoid invaded by the four bacteria strains back to the mouse through tail vein and analyzed the metastasis colonization using tissue clearing imaging. Indeed, *Staphylococcus* and *Lactobacillus* significantly increased the number of colonized metastatic tumor foci, while *Enterococcus* and *Streptococcus* had minimal effect (Figures 5E–5G), possibly due to the low bacteria-invasion efficiency at individual cell or organoid level.

To study whether these bacteria strains can trigger metastasis at the primary tumor site, we reintroduced various bacteria strains back into the antibiotics-conditioned primary tumor (Shi et al., 2020). The occupation of the injected bacteria in the tumor were validated by culturing tumor homogenates (Figure S5C). The intratumor re-administration of various strains of bacteria (including *S. xylosus*, *L. animalis*, and *S. cuniculi*) left the primary tumor growth unchanged but triggered an increase of the lung metastasis compared to Dox-treated bacteria-free PyMT tumors

### Figure 4. PyMT tumor cells carry over microbiota during metastasis

- (A) Schematic diagram showing the sample information and the abbreviations.  
(B) Stacked bar plot showing the bacteria relative abundance at the phylum and genus level in various samples.  
(C) Supervised unifracs beta diversity analysis of various samples. normal breast (NB).  
(D) Microbiota correlation analysis by unifracs distance.  
(E) Bacteria property analysis by bugbase.  
(F) 16S FISH analysis of lung section with PyMT tumor metastasis. Yellow arrow pointing to 16S positive bacteria, dotted lines indicating metastasis foci. Data are presented as mean  $\pm$  SD. Red, 16S probe; Blue, DAPI. Scale bar, 100 $\mu$ m.  
(G) Staining and quantification of bacteria in circulating tumor cells with Keratin 8 + 18 antibody and 16S probe in blood. Green, Keratin 8 + 18, Red, 16S probe. Scale bars, 50 $\mu$ m, 20 $\mu$ m.  
(H) Heatmap showing the dynamic distribution of various bacteria strains during tumor progression.  
(I) The dynamics of relative abundance of four distinct bacteria clusters within various samples.  
(J) Schematic diagram showing an *erm*-GFP cassette was inserted at *Lldh* genome locus of *S. xylosus*. Right: different strategies to introduce bacteria into mice via intratumor or tail vein injection.  
(K) Example of positive and negative recombinant *S. xylosus* clones judged by PCR.  
(L) Analysis of positive clones for recombinant *S. xylosus* in the lung under various bacteria administration strategies.  $N_{\text{tumor inj}}^{\text{met}+}=16$ ,  $N_{\text{tumor inj}}^{\text{met}-}=3$ ,  $N_{\text{Bac tail vein}}=5$ ,  $N_{\text{Bac-tumor tail vein}}=5$ .  
See also Figure S4 and Table S3.



(legend on next page)

(Figures 5H–5J). *E. faecalis* is an exception in promoting lung metastasis, possibility due to its low invasion efficiency. This argues that invasion into the tumor cells might be required to exert metastasis-stimulating function. Intriguingly, when we performed this experiment on an MMTV-Wnt tumor mouse model, which is known to rarely metastasize (Li et al., 2000) and contained low abundance microbiota (Figures S5D–S5F), the tumor started to metastasize after one shot of the on-site administration of bacteria, reinforcing the role of certain bacteria in promoting metastasis initiation (Figures 5K–5M). Of note, this metastasis promoting effect was not dependent on the innate immune activation triggered by bacteria injection (Figures S5G and S5H), because enforced bacteria invasion via cell impermeable antibiotics (Kalaora et al., 2021) also enhanced lung metastasis with unchanged primary tumors and unperturbed immunity (Figures 5N–5R). These coherent *in vitro* and *in vivo* functional assays demonstrate that the invasion of bacteria can functionally promote the BT metastasis in different tumor genetic mouse models.

### Intracellular bacteria enhance the viability of tumor cells under mechanical stress

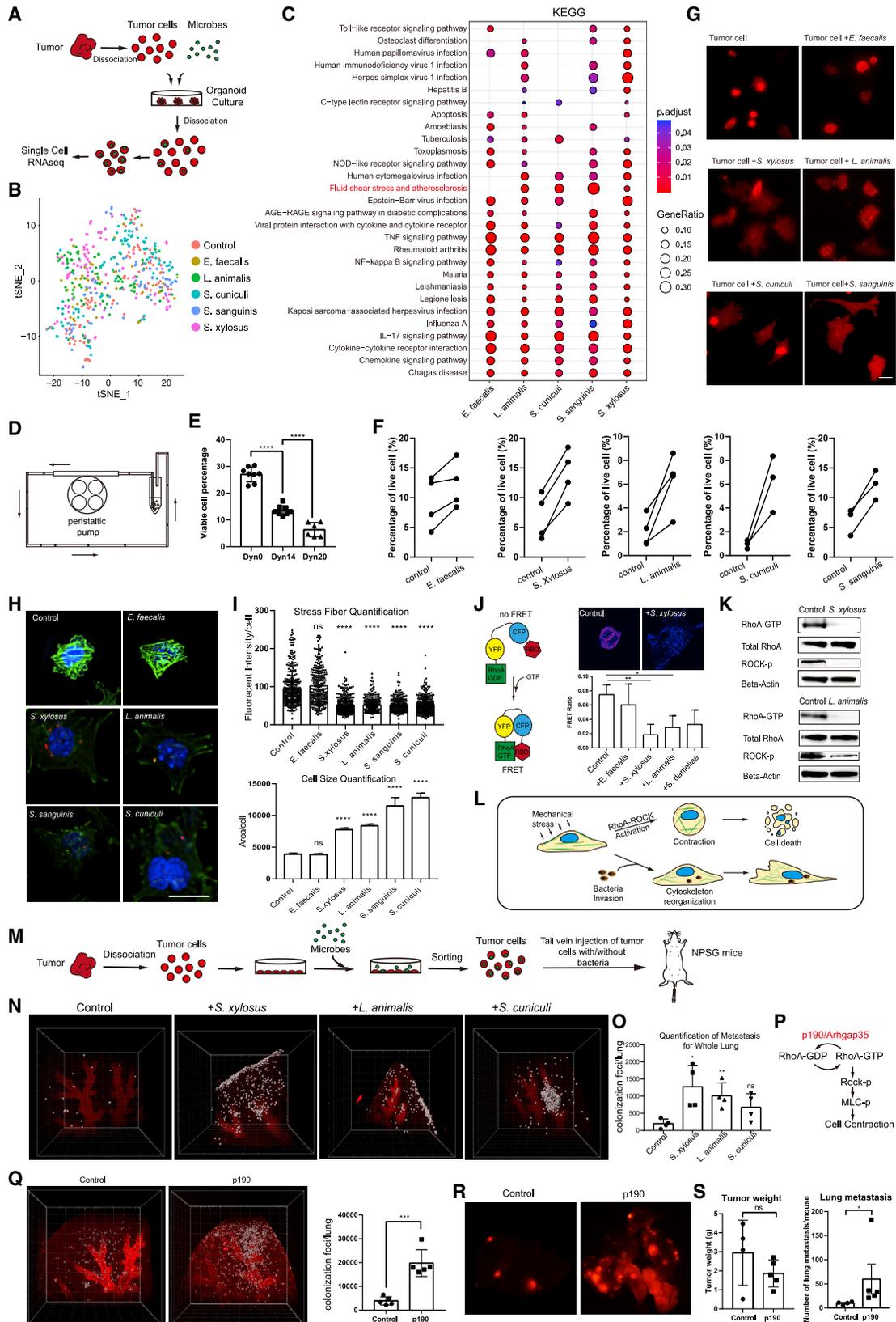
We next seek to understand what microbes do within the cytoplasm and how they contribute to the cancer cell invasiveness. We performed single-cell RNAseq analysis of bacteria invaded cancer cells isolated from organoid culture (Figure 6A). Overall, we extracted 5,023 genes within 483 cells in total. It appears that bacteria infected Krt14, Krt8, and Krt18 tumor cells equally (Figure S6A). t-distributed stochastic neighbor embedding (tSNE) plot showed that the cancer cell invasion of various bacteria did not change the cancer cell heterogeneity (Figure 6B), but enhanced stem cell activity (Spike et al., 2012) (Figure S6D). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis revealed that tumor cells invaded by bacteria turned on many immunity-related signals, including the Toll-like receptor signaling pathway, the NOD-like receptor signaling pathway, the TNF signaling pathway, the NF- $\kappa$ B signaling pathway, the Th17 signaling pathway, and the cytokine chemokine related pathways etc. (Figure 6C; Table S4). Interestingly, we noticed that the invasion of *S. xyloso*, *L. animalis*, *S. cuniculi*, and *S. sanguinis* into tumor cells specifically triggered the fluid shear stress pathway, but the low invasion strain *E. faecalis* did not (Figures 6C and S6B–S6C; Table S4).

For the metastasizing cancer cells, they experience fluid shear stress particularly after intravasation into the circulation system, which frequently triggers cellular apoptosis (Mitchell and King, 2013). The involvement of bacteria into the fluid shear stress pathway indicates that bacteria may modulate the stress response and influence the cancer cell viability. To test this idea, we set up a circulation system with peristaltic pump (Regmi et al., 2017) to mimic the fluid shear stress in the blood vessel (Figure 6D), in which tumor cells showed gradually decreased cell viability as the increase of stress (Figure 6E) within the physiological range (Follain et al., 2020). Indeed, after the median *in vitro* fluid stress, cells with bacteria showed a higher survival rate than cells without bacteria (Figures 6F and 6G). Myeloid-derived suppressor cells (MDSCs) induced by intratumor bacteria injection were not involved in regulating this mechanical resistance (Figures S7E and S7F). Importantly, the live cells with bacteria adhered and spread better on the plate and showed much larger size than control, indicating an alteration of cytoskeleton and the attachment ability (Figures 6H and 6I). This phenotype indicated a possible role for bacteria in the organization of the actin cytoskeleton. Consistent with this notion, phalloidin staining of the cells showed a significantly reduced stress fiber intensity after the invasion of bacteria *in vitro* in the culture dish (Figures 6H and 6I) and *in vivo* for the primary cells (Figures S7A–S7C), suggesting that the mechanical stress induced contractile forces can be relieved by the cytoplasmic bacteria (Figure 6L).

For actin cytoskeleton organization in cells, Rho GTPase family, especially the well-characterized members RhoA, Rac, and Cdc42, are responsible for regulating stress fiber, lamellipodia, and filopodia, respectively (Nobes and Hall, 1995). In addition, in the stem cell isolation and culture assay, ROCK kinase (downstream of RhoA) inhibitor is frequently utilized to prevent mechanical force triggered apoptosis (Ohgushi et al., 2010). ROCK inhibitor does this by the inhibition of phosphorylation of MLC and the disassembly of stress fiber, relieving contractile forces (Shi et al., 2013). We speculated that the intracellular bacteria may play a similar role. This idea was supported by the fluorescence resonance energy transfer (FRET) sensor analysis of activated RhoA and Western blot of RhoA-GTP, showing the invasion of bacteria indeed suppressed RhoA and ROCK activation (Figures 6J, 6K, S7G, and S7H), and inhibition of ROCK abolished viability difference caused by bacteria invasion (Figure S7D).

### Figure 5. Tumor-resident microbiota promote metastasis

- (A) 3D co-culture experiment showing CFSE labeled bacteria invaded PyMT mTmG tumor organoids. Red, tdTomato; Green, bacteria. Scale bar, 10 $\mu$ m.  
(B) An optical section (left) and 3D reconstructed view (middle and right) of mTmG tumor organoid containing *S. xyloso*. Red, tdTomato; Green, *S. xyloso*; Blue, DAPI. Scale bar, 10 $\mu$ m.  
(C) EM image showing intracellular bacteria in the tumor organoid. Yellow arrow heads pointing to the cleavage furrow of a dividing *Staphylococcus*.  
(D) Bacteria invasion efficiency quantified by FACS of dissociated individual tumor cells. Data are presented as mean  $\pm$  SD.  
(E) Top: schematic diagram showing the workflow of metastatic colonization assay with organoids. Bottom: image of the cleared lung from the mice IV injected with control (left) or *S. xyloso* invaded (right) tumor organoids. Yellow arrow heads pointing to the organoid foci in the lung.  
(F) The ratio of bacteria positive organoids. Data are presented as mean  $\pm$  SD.  
(G) Organoid foci in the lung after tail vein injection of bacteria-containing organoids. Data are presented as mean  $\pm$  SD.  
(H and K) Schematic diagram showing the workflow of intratumor injection of bacteria to MMTV-PyMT mTmG tumor (H) or MMTV-Wnt tumor (K).  
(I, J, L, and M) Total PyMT spontaneous tumor and Wnt tumor weight (I and L) and metastasis burden (J and M) after intratumor injection of bacteria. Data are presented as mean  $\pm$  SEM. Mann Whitney test P value. \*  $p < 0.05$ , \*\*  $p < 0.01$ .  
(N) Schematic diagram showing IV administration of ampicillin + gentamicin on PyMT-mTmG tumor-bearing NPSG mice.  
(O–R) Characterization of the intracellular bacteria abundance, primary tumor weight and metastasis after ampicillin + gentamicin treatment.  
See also Figure S5.



(legend on next page)

To test whether the bacteria-induced relief of contractile forces benefits circulating tumor cells' survival under blood stream shear forces *in vivo*, we injected bacteria associated cancer cells intravenously back to the immunodeficient NPSG mice via tail vein to avoid the indirect influence from immune system (Figure 6M). We found that *S. xylosum*, *L. animalis*, and *S. cuniculi* can significantly increase the survival of cancer cells in the lung by 6.4-fold, 5.0-fold, and 3.4-fold, respectively (Figures 6N and 6O). The influence of *E. faecalis* could not be tested in this assay due to its low invasion efficiency. When we directly inhibited the downstream RhoA signaling by enforced expression of p190/Arhgap35 that hydrolyze GTP (Holinstat et al., 2006) (Figure 6P), the stress fiber formation *in vitro* was drastically suppressed (Figures S7I and S7J), and the tumor cell's colonization in the lung was strikingly enhanced (Figures 6P and 6Q). In addition, when we grafted control and p190 expressing cells onto immunocompromised NPSG mice, we observed a similar primary tumor weight but a prominent increase of lung metastasis for p190 expressing tumors (Figures 6R and 6S), indicating the intrinsic signaling modulation was sufficient to drive metastatic colonization. These data suggest that the invasion of *S. xylosum*, *L. animalis*, and *S. cuniculi* enhances cancer cells colonization in the lung *in vivo* after stressed by the fluid shear stress in the circulation system. This explains why the antibiotic elimination of tumor-resident microbiota impeded metastasis development.

### Conserved microbiota profile in human BT

The microbiota within human BTs has been characterized by 16S sequencing and pan-pathogen microarray (Pathochip) (Banerjee et al., 2018; Banerjee et al., 2015; Costantini et al., 2018; Hieken et al., 2016; Urbaniak et al., 2016; Xuan et al., 2014). But the microbiota link between the primary tumor and the metastatic sites is not clear in human breast cancer. We collected human BT tissue with paired tumor adjacent normal breast tissue and the lymph node metastasis, and analyzed the composition of microbiota by 16S sequencing. The bacteria loads within BT tissue, tumor adjacent breast tissue, and lymph node metastasis are significantly higher than the environment control, reaching an

average of  $10^5$  equivalent bacteria/gram of tissue, similar to what we detected in mouse BT tissue (Figure 7A). In addition, consistent with mouse BT, the normal human breast and the human BT have significantly increased Firmicutes compared with EBC (Figures 7B, 7C, S7K, and S7L). Moreover, in line with the mouse data, human tumor microbiota and lymph node metastasis microbiota had distinct microbial community, significantly reduced alpha diversity, and increased facultative anaerobes compared with human normal breast samples (Figures 7D–7F) ( $p < 0.05$ ). Interestingly, the lymph node metastasis is closely clustered with human BT microbiota, supporting the notion that the microbes in the metastasis are inherited from the primary BT. The BT tissue contains significant higher abundance of *Enterococcus* and *Streptococcus* (Figure 7H; Table S5), reminiscent of that in mouse tumor data. These data collectively show that murine BT and human BT have a similar microbial community profile and dynamics, suggesting that the microbiota in human BT may play a conserved role in the human cancer pathogenesis and progression.

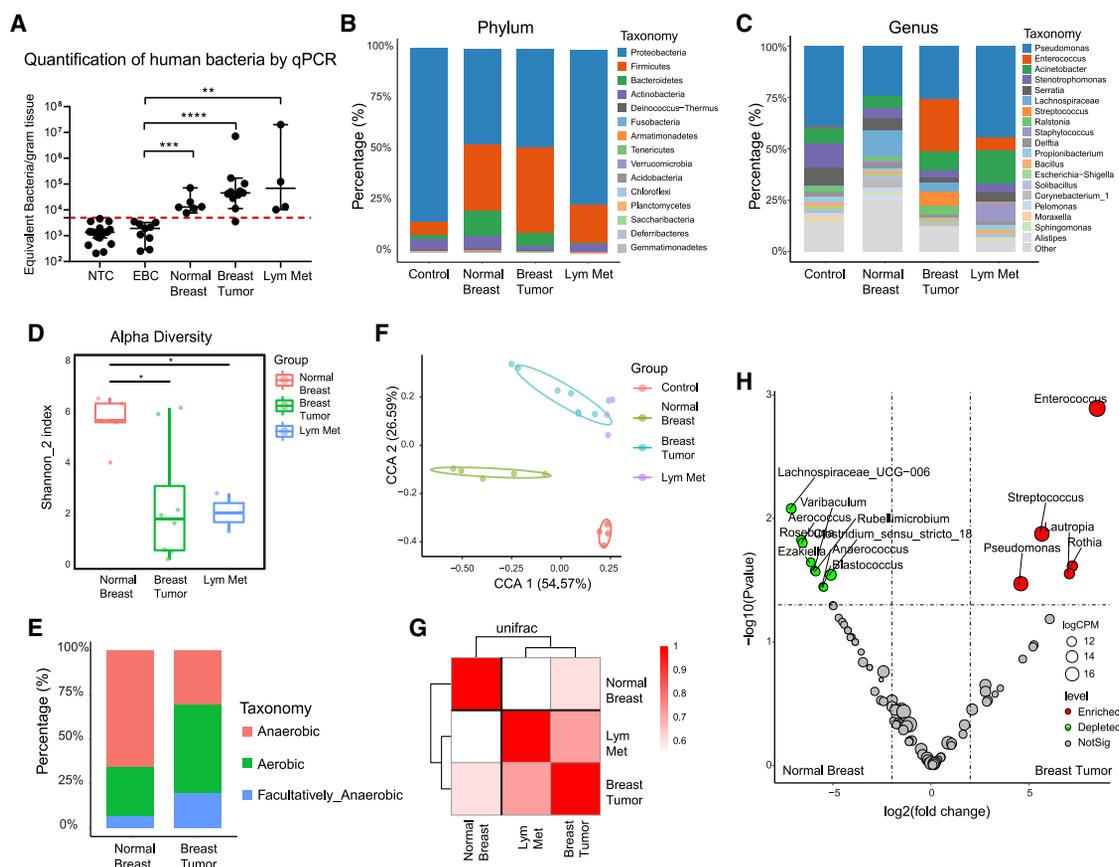
### DISCUSSION

Recently, there has been increasing awareness that human tumors contain significant amount of viable commensal microbiota (Banerjee et al., 2018; Banerjee et al., 2015; Costantini et al., 2018; Hieken et al., 2016; Nejman et al., 2020; Riquelme et al., 2019; Urbaniak et al., 2016; Xuan et al., 2014). Whether these microbes are passengers or drivers of tumor progression is an intriguing question that emerges. In the present study, using a spontaneous mouse BT model, we characterized the BT microbiota profile and developed methods to dissect the different roles of BT-resident microbiota and gut-resident microbiota. We showed that the intracellular microbiota plays crucial roles in tumor metastasis by modulating cellular cytoskeleton and cell viability upon mechanical stress. For the vast majority of the murine PyMT tumor samples and human breast cancer patients, we have successfully detected the presence of microbiota; this indicates that the intratumor microbiota is more likely to be an intrinsic and integral component of the tumor tissue instead of

### Figure 6. Intracellular microbiota reorganizes cytoskeleton and promotes resistance against mechanical stress

- (A) Schematic diagram showing the workflow of single cell RNAseq for bacteria-containing PyMT tumor cells.  
(B and C) t-SNE plot and bubble plot analysis showing the population and enriched pathways for bacteria-containing PyMT tumor cells.  
(D) Schematic diagram showing the *in vitro* setup of fluid shear stress model by peristaltic pump mimicking blood flow.  
(E) Tumor cell viability under various levels of fluid shear stress at 0, 14, 20 Dyn  $\text{cm}^{-2}$ . Unpaired t test, \*\*\*\*  $p < 0.0001$ .  
(F and G) Images and viability analysis of tumor cells w or w/o bacteria under fluid shear stress at 14 Dyn  $\text{cm}^{-2}$ .  
(H and I) Phalloidin staining and quantification of re-plated PyMT tumor cells w or w/o invaded bacteria, showing the actin stress fiber. Green, Phalloidin, Red, Vancomycin probe detecting gram positive bacteria. Scale bar, 10  $\mu\text{m}$ .  $n = 3$ . Welch's t test P value.  
(J) RhoA-GTP activity quantification by FRET sensor. Quantification of FRET ratios were plotted for each individual bacteria strain.  
(K) Western blot showing the levels of RhoAGTP, ROCK2-p after bacteria invasion.  
(L) Schematic diagram showing the consequence of fluid shear stress in the presence or absence of bacteria, or ROCK inhibitor Y27632.  
(M) Schematic diagram showing the workflow of the *in vivo* survival assay for tail vein injection of tumor cells with or without bacteria.  
(N and O) Images and quantification of cleared lung tissue for the survival metastatic tumor cells with various intracellular bacteria. Unpaired t test, \*  $p < 0.05$ , \*\*  $p < 0.01$ .  
(P) Schematic diagram showing the regulation of RhoA's inactivation by p190/Arhgap35.  
(Q) Metastatic colonization capacity of control and p190 expressing PyMT tumor cells after IV injection. Unpaired t test \*\*\*  $p < 0.001$   
(R and S) Lung Metastasis analysis of orthotopic grafting of control or p190 expressing tumor cells. Unpaired t test (for tumor weight) and Welch's test (for metastasis) (data are presented as mean  $\pm$  SEM) was used to calculate the two tailed P value. \*  $p < 0.05$ . Data in this figure are all presented as mean  $\pm$  SD unless indicated.

See also Figures S6 and S7 and Table S4.



**Figure 7. Human breast tumor microbiota showed correlation between primary site and lymph node metastasis**

(A) Quantification of bacteria load in paired normal human breast tissue, human breast tumor and the lymph node metastasis. Data are presented as median  $\pm$  95% CI. Mann Whitney test P value. \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ .

(B and C) Stacked bar plot of relative abundance of bacteria at the phylum and genus level in EBC control, human normal breast tissue, breast tumor, and lymph node metastasis.

(D) Shannon index of human normal breast tissue, breast tumor, and lymph node metastasis. \*  $p < 0.05$ .

(E) Bar plot showing the composition of aerobics, anaerobics, and facultative anaerobics in human normal breast tissue and human breast tumor.

(F) Supervised clustering of tissue samples using unifrac distance.

(G) Correlation analysis of the bacteria communities using unifrac distance.

(H) Volcano plot analysis of the microbiota in normal and breast tumor.

See also [Figure S7](#) and [Table S5](#).

an incidental presence due to pathogenic infections. This notion is also supported by the recent comprehensive intratumor microbiota analysis in various human cancer types (Nejman et al., 2020). Therefore, the tumor cells hijacked by microbes could be more common than so far known in cancer patients, which underscores the broad clinical value of understanding the exact role of the tumor-resident microbial community in cancer progression.

Metastatic colonization has been reported to be a highly inefficient process with drastic tumor cell death when reaching the distal organs (Massagué and Obenauf, 2016). The colonization efficiency is therefore the bottleneck for successful metastasis. Any promotion of the cell viability via genetic alterations (e.g., mutations) (Bos et al., 2009; Ishaque et al., 2018; Minn et al., 2005), epigenetic states (e.g., epithelial mesenchymal transition [EMT]) (Ocaña et al., 2012; Tsai et al., 2012), or microenvironment factors (e.g., cancer-activated fibroblasts and immune cells) (Massagué and Obenauf, 2016), can lead to the survival of the deadly tumor

cells that seed and initiate metastatic foci and finally become catastrophic to the patient. However, most current chemo drugs often target tumor cell growth rather than tumor cell seeding. Our findings provide a strategy that cancer cells utilize intracellular microbiota to survive the fluid shear stress in the circulation system during metastatic colonization without affecting tumor growth. As the fluid shear stress was estimated to be low in the tissue interstitial space ( $0.1-1 \text{ Dyn cm}^{-2}$ ) and highest in the artery ( $4-30 \text{ Dyn cm}^{-2}$ ) (Follain et al., 2020), the survival benefits of tumor bacteria on tumor cells are most prominent during metastasis rather than primary tumor growth. This mechanism is not restricted to breast cancer, because there is also evidence in colorectal cancer that intratumor microbiota can persist during metastasis and passages (Bullman et al., 2017). Therefore, the intracellular microbiota could be a potential target for preventing metastasis in broad cancer types at an early stage, which is much better than to have to treat it later on.

On the clinical side, the intracellular microbiota may lead to distinct biological consequences from extracellular microbiota upon the invasion of tumor cells. These intracellular microbes would also respond differently than extracellular counterparts to various kinds of antibiotics. In addition, the gut microbiome has been implicated in modulating cancer immune response (Gopalakrishnan et al., 2018; Sivan et al., 2015; Vétizou et al., 2015), adding another layer of complexity. This may partially explain why the clinical use of antibiotics in cancer patients ended up with disparate results in cancer progression when combined with different therapeutics (Derosa et al., 2018; Kim et al., 2019; Lindahl et al., 2019; Pinato et al., 2019; Routy et al., 2018; Scatena et al., 2018). Various types of antibiotics via different administration routes can cause various consequences against gut microbiota, extracellular tumor microbiota, and intracellular tumor microbiota, leading to intricate outcomes. In the future, further in-depth analysis of how the bacteria invade into tumor cells, how the intracellular bacteria are integrated into the host-cell system, and how the bacteria-containing tumor cells interact with the immune system will provide us insights on how to properly implement antibiotics for cancer therapeutics in the clinic.

### Limitations of the study

Although our study revealed a clear role of intratumor bacteria in promoting cancer cell metastatic colonization, it remains possible that the gut microbiome and immune system may act together with intratumor bacteria in determining cancer progression. It would also be interesting to explore whether intratumor bacteria exert certain roles during tumor cell dissemination, intravasation, and extravasation as well as cancer dormancy. More specifically, the innate immunity triggered by intratumor bacteria injection could possibly be involved in certain steps in bacteria driven metastasis other than regulating mechanical property, which remains an open question.

### 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 SUBJECT DETAILS
  - Mice
  - Human samples collection
- METHOD DETAILS
  - Mouse mammary tumor dissection and dissociation
  - Bacteria culture and identification
  - Quantification and profiling of tissue-resident microbiome
  - Microbiome data analysis
  - Antibiotics treatments on MMTV-PyMT tumor cells
  - Antibiotics treatments on MMTV-PyMT mouse model
  - Antibiotics Treatment on Germ Free Mouse

- Injection of intratumor bacteria to MMTV-PyMT and MMTV-Wnt mTmG tumor
- 16S RNA FISH Assay
- Immunofluorescence assays
- Immunohistochemistry assays
- Sample Preparation for Transmission Electron Microscopy
- Flow Cytometric Enumeration of Faecal Bacterial Loads
- Flow Cytometry Profiling of Tumor Immune Cells
- Circulating Tumor Cell Analysis
- Genetic Labeling of *Staphylococcus xylosus*
- Bacteria Tracing Experiment
- Cell viability test and F-actin staining
- Microfluidic Circulatory System
- Rho activity measurement by FRET
- Western Blot
- 3D co-culture system and organoid imaging
- Tissue clearing, imaging and analysis
- Single cell sequencing
- Single cell RNA-seq data analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cell.2022.02.027>.

### ACKNOWLEDGMENTS

We thank Drs. Yalin Wang, Lei Gao, Fang Xiao, and Yinping Gao at the Microscopy Core Facility of Westlake University for advice and assistance in electron microscopy sample preparation and data collection. We thank the Flow Cytometry Facility of Westlake University for assistance in FACS analysis and sorting. We thank the Supercomputer Center of Westlake University for microbiome analysis. We thank the Westlake Animal Facility for mouse husbandry. We thank Dr. Wei Wang for the generous gift of vancomycin probe. We are grateful for Dr. Dangsheng Li, Dr. Hongtao Yu, Dr. Jinyi Xiang, Dr. Liang Tao, and Dr. Feng Ju for valuable discussions and suggestions. We appreciate help from Linda Jacqueline van Weele in revising the manuscript. This work was supported by National Natural Science Foundation of China (NSFC) grants 32170803 and 81872405. This work was supported by Westlake Education Foundation.

### AUTHOR CONTRIBUTIONS

Conceptualization, S.C., A.F., B.Y., and T.D.; Methodology, S.C., A.F., B.Y., T.D., and C.W.; Investigation, A.F., B.Y., T.D., H.B., Y.Z., and Y.G.; Formal Analysis, X.L., H.L., and N.L. (specifically, X.L. analyzed single cell RNAseq data, and H.L. and N.L. analyzed microbiome data); Resources, Y.C., J.Y. and Y.L.; Writing – Original Draft, S.C., A.F., B.Y., and T.D.; Writing – Review & Editing, Y.C. and N.L.; Supervision, S.C.

### DECLARATION OF INTERESTS

All authors declare no competing interests.

Received: December 1, 2020  
 Revised: January 10, 2022  
 Accepted: February 24, 2022  
 Published: April 7, 2022

## REFERENCES

- Aceto, N., Bardia, A., Miyamoto, D.T., Donaldson, M.C., Wittner, B.S., Spencer, J.A., Yu, M., Pely, A., Engstrom, A., Zhu, H., et al. (2014). Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell* 158, 1110–1122.
- Banerjee, S., Wei, Z., Tan, F., Peck, K.N., Shih, N., Feldman, M., Rebbeck, T.R., Alwine, J.C., and Robertson, E.S. (2015). Distinct microbiological signatures associated with triple negative breast cancer. *Sci. Rep.* 5, 15162.
- Banerjee, S., Tian, T., Wei, Z., Shih, N., Feldman, M.D., Peck, K.N., DeMichele, A.M., Alwine, J.C., and Robertson, E.S. (2018). Distinct Microbial Signatures Associated With Different Breast Cancer Types. *Front. Microbiol.* 9, 951.
- Bos, P.D., Zhang, X.H., Nadal, C., Shu, W., Gomis, R.R., Nguyen, D.X., Minn, A.J., van de Vijver, M.J., Gerald, W.L., Foekens, J.A., and Massagué, J. (2009). Genes that mediate breast cancer metastasis to the brain. *Nature* 459, 1005–1009.
- Brückner, R. (1997). Gene replacement in *Staphylococcus carnosus* and *Staphylococcus xylosum*. *FEMS Microbiol. Lett.* 151, 1–8.
- Buchta Rosean, C., Bostic, R.R., Ferey, J.C.M., Feng, T.Y., Azar, F.N., Tung, K.S., Dozmorov, M.G., Smirnova, E., Bos, P.D., and Rutkowski, M.R. (2019). Preexisting Commensal Dysbiosis Is a Host-Intrinsic Regulator of Tissue Inflammation and Tumor Cell Dissemination in Hormone Receptor-Positive Breast Cancer. *Cancer Res.* 79, 3662–3675.
- Bullman, S., Pedamallu, C.S., Sicinska, E., Clancy, T.E., Zhang, X., Cai, D., Neuberger, D., Huang, K., Guevara, F., Nelson, T., et al. (2017). Analysis of *Fusobacterium* persistence and antibiotic response in colorectal cancer. *Science* 358, 1443–1448.
- Butler, A., Hoffman, P., Smibert, P., Papalexis, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* 36, 411–420.
- Cai, S., Kalisky, T., Sahoo, D., Dalerba, P., Feng, W., Lin, Y., Qian, D., Kong, A., Yu, J., Wang, F., et al. (2017). A Quiescent Bcl11b High Stem Cell Population Is Required for Maintenance of the Mammary Gland. *Cell Stem Cell* 20, 247–260 e245.
- Cheung, K.J., Padmanaban, V., Silvestri, V., Schipper, K., Cohen, J.D., Fairchild, A.N., Gorin, M.A., Verdone, J.E., Pienta, K.J., Bader, J.S., and Ewald, A.J. (2016). Polyclonal breast cancer metastases arise from collective dissemination of keratin 14-expressing tumor cell clusters. *Proc. Natl. Acad. Sci. USA* 113, E854–E863.
- Costantini, L., Magno, S., Albanese, D., Donati, C., Molinari, R., Filippone, A., Masetti, R., and Merendino, N. (2018). Characterization of human breast tissue microbiota from core needle biopsies through the analysis of multi hypervariable 16S-rRNA gene regions. *Sci. Rep.* 8, 16893.
- Dapito, D.H., Mencin, A., Gwak, G.Y., Pradere, J.P., Jang, M.K., Mederacke, I., Caviglia, J.M., Khiabanian, H., Adeyemi, A., Bataller, R., et al. (2012). Promotion of hepatocellular carcinoma by the intestinal microbiota and TLR4. *Cancer Cell* 21, 504–516.
- Davis, N.M., Proctor, D.M., Holmes, S.P., Relman, D.A., and Callahan, B.J. (2018). Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* 6, 226.
- de Goffau, M.C., Lager, S., Salter, S.J., Wagner, J., Kronbichler, A., Charnock-Jones, D.S., Peacock, S.J., Smith, G.C.S., and Parkhill, J. (2018). Recognizing the reagent microbiome. *Nat. Microbiol.* 3, 851–853.
- Derosa, L., Hellmann, M.D., Spaziano, M., Halpenny, D., Fidelle, M., Rizvi, H., Long, N., Plodkowski, A.J., Arbour, K.C., Chaff, J.E., et al. (2018). Negative association of antibiotics on clinical activity of immune checkpoint inhibitors in patients with advanced renal cell and non-small-cell lung cancer. *Ann. Oncol.* 29, 1437–1444.
- DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P., and Andersen, G.L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* 72, 5069–5072.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21.
- Dong, J., Hu, Y., Fan, X., Wu, X., Mao, Y., Hu, B., Guo, H., Wen, L., and Tang, F. (2018). Single-cell RNA-seq analysis unveils a prevalent epithelial/mesenchymal hybrid state during mouse organogenesis. *Genome Biol.* 19, 31.
- Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461.
- Edgar, R.C. (2016). SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS sequences. Preprint at bioRxiv. <https://doi.org/10.1101/074161>.
- Eisenhofer, R., Minich, J.J., Marotz, C., Cooper, A., Knight, R., and Weyrich, L.S. (2019). Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. *Trends Microbiol.* 27, 105–117.
- Flemer, B., Lynch, D.B., Brown, J.M., Jeffery, I.B., Ryan, F.J., Claesson, M.J., O'Riordain, M., Shanahan, F., and O'Toole, P.W. (2017). Tumour-associated and non-tumour-associated microbiota in colorectal cancer. *Gut* 66, 633–643.
- Follain, G., Herrmann, D., Harlepp, S., Hyenne, V., Osmani, N., Warren, S.C., Timpson, P., and Goetz, J.G. (2020). Fluids and their mechanics in tumour transit: shaping metastasis. *Nat. Rev. Cancer* 20, 107–124.
- Garrett, W.S. (2015). Cancer and the microbiota. *Science* 348, 80–86.
- Geller, L.T., and Straussman, R. (2017). Intratumoral bacteria may elicit chemoresistance by metabolizing anticancer agents. *Mol. Cell. Oncol.* 5, e1405139.
- Geller, L.T., Barzily-Rokni, M., Danino, T., Jonas, O.H., Shental, N., Nejman, D., Gavert, N., Zwang, Y., Cooper, Z.A., Shee, K., et al. (2017). Potential role of intratumor bacteria in mediating tumor resistance to the chemotherapeutic drug gemcitabine. *Science* 357, 1156–1160.
- Gopalakrishnan, V., Spencer, C.N., Nezi, L., Reuben, A., Andrews, M.C., Karpman, T.V., Prieto, P.A., Vicente, D., Hoffman, K., Wei, S.C., et al. (2018). Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science* 359, 97–103.
- Guy, C.T., Cardiff, R.D., and Muller, W.J. (1992). Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. *Mol. Cell. Biol.* 12, 954–961.
- Hieken, T.J., Chen, J., Hoskin, T.L., Walther-Antonio, M., Johnson, S., Ramaker, S., Xiao, J., Radisky, D.C., Knutson, K.L., Kalari, K.R., et al. (2016). The Microbiome of Aseptically Collected Human Breast Tissue in Benign and Malignant Disease. *Sci. Rep.* 6, 30751.
- Holinstat, M., Knezevic, N., Broman, M., Samarel, A.M., Malik, A.B., and Mehta, D. (2006). Suppression of RhoA activity by focal adhesion kinase-induced activation of p190RhoGAP: role in regulation of endothelial permeability. *J. Biol. Chem.* 281, 2296–2305.
- Ibekwe, A.M., Watt, P.M., Grieve, C.M., Sharma, V.K., and Lyons, S.R. (2002). Multiplex fluorogenic real-time PCR for detection and quantification of *Escherichia coli* O157:H7 in dairy wastewater wetlands. *Appl. Environ. Microbiol.* 68, 4853–4862.
- Iida, N., Dzutsev, A., Stewart, C.A., Smith, L., Bouladoux, N., Weingarten, R.A., Molina, D.A., Salcedo, R., Back, T., Cramer, S., et al. (2013). Commensal bacteria control cancer response to therapy by modulating the tumor microenvironment. *Science* 342, 967–970.
- Ishaque, N., Abba, M.L., Hauser, C., Patil, N., Paramasivam, N., Huebschmann, D., Leupold, J.H., Balasubramanian, G.P., Kleinheinz, K., Toprak, U.H., et al. (2018). Whole genome sequencing puts forward hypotheses on metastasis evolution and therapy in colorectal cancer. *Nat. Commun.* 9, 4782.
- Jervis-Bardy, J., Leong, L.E., Marri, S., Smith, R.J., Choo, J.M., Smith-Vaughan, H.C., Nosworthy, E., Morris, P.S., O'Leary, S., Rogers, G.B., and Marsh, R.L. (2015). Deriving accurate microbiota profiles from human samples with low bacterial content through post-sequencing processing of Illumina MiSeq data. *Microbiome* 3, 19.
- Jin, C., Lagoudas, G.K., Zhao, C., Bullman, S., Bhutkar, A., Hu, B., Arneh, S., Sandel, D., Liang, X.S., Mazzilli, S., et al. (2019). Commensal Microbiota Promote Lung Cancer Development via gamma delta T Cells. *Cell* 176, 998–1013 e1016.

- Kalaora, S., Nagler, A., Nejman, D., Alon, M., Barbolin, C., Barnea, E., Kete-laars, S.L.C., Cheng, K., Vervier, K., Shental, N., et al. (2021). Identification of bacteria-derived HLA-bound peptides in melanoma. *Nature* **592**, 138–143.
- Kim, D., Hofstaedter, C.E., Zhao, C., Mattei, L., Tanes, C., Clarke, E., Lauder, A., Sherrill-Mix, S., Chehoud, C., Kelsen, J., et al. (2017). Optimizing methods and dodging pitfalls in microbiome research. *Microbiome* **5**, 52.
- Kim, H., Lee, J.E., Hong, S.H., Lee, M.A., Kang, J.H., and Kim, I.H. (2019). The effect of antibiotics on the clinical outcomes of patients with solid cancers undergoing immune checkpoint inhibitor treatment: a retrospective study. *BMC Cancer* **19**, 1100.
- Kolde, R. (2019). pheatmap: Pretty Heatmaps. R package version 1.0.12. <https://CRAN.R-project.org/package=pheatmap>.
- Kumar, R., Herold, J.L., Schady, D., Davis, J., Kopetz, S., Martinez-Moczygemba, M., Murray, B.E., Han, F., Li, Y., Callaway, E., et al. (2017). *Streptococcus gallolyticus* subsp. *gallolyticus* promotes colorectal tumor development. *PLoS Pathog.* **13**, e1006440.
- Laurence, M., Hatzis, C., and Brash, D.E. (2014). Common contaminants in next-generation sequencing that hinder discovery of low-abundance microbes. *PLoS ONE* **9**, e97876.
- Lehar, S.M., Pillow, T., Xu, M., Staben, L., Kajihara, K.K., Vandlen, R., DePalatis, L., Raab, H., Hazenbos, W.L., Morisaki, J.H., et al. (2015). Novel antibody-antibiotic conjugate eliminates intracellular *S. aureus*. *Nature* **527**, 323–328.
- Li, Y., Hively, W.P., and Varmus, H.E. (2000). Use of MMTV-Wnt-1 transgenic mice for studying the genetic basis of breast cancer. *Oncogene* **19**, 1002–1009.
- Li, J., Sung, C.Y., Lee, N., Ni, Y., Pihlajamäki, J., Panagiotou, G., and El-Nezami, H. (2016). Probiotics modulated gut microbiota suppresses hepatocellular carcinoma growth in mice. *Proc. Natl. Acad. Sci. USA* **113**, E1306–E1315.
- Li, Y., Tinoco, R., Elmen, L., Segota, I., Xian, Y., Fujita, Y., Sahu, A., Zarecki, R., Marie, K., Feng, Y., et al. (2019). Gut microbiota dependent anti-tumor immunity restricts melanoma growth in *Rnf5*<sup>-/-</sup> mice. *Nat. Commun.* **10**, 1492.
- Lindahl, L.M., Willerslev-Olsen, A., Gjerdrum, L.M.R., Nielsen, P.R., Blümel, E., Rittig, A.H., Celis, P., Herpers, B., Becker, J.C., Stausbøl-Grøn, B., et al. (2019). Antibiotics inhibit tumor and disease activity in cutaneous T-cell lymphoma. *Blood* **134**, 1072–1083.
- Liu, Y.X., Qin, Y., Chen, T., Lu, M., Qian, X., Guo, X., and Bai, Y. (2021). A practical guide to amplicon and metagenomic analysis of microbiome data. *Protein Cell* **12**, 315–330.
- Massagué, J., and Obenauf, A.C. (2016). Metastatic colonization by circulating tumour cells. *Nature* **529**, 298–306.
- Minn, A.J., Gupta, G.P., Siegel, P.M., Bos, P.D., Shu, W., Giri, D.D., Viale, A., Olshen, A.B., Gerald, W.L., and Massagué, J. (2005). Genes that mediate breast cancer metastasis to lung. *Nature* **436**, 518–524.
- Mitchell, M.J., and King, M.R. (2013). Fluid Shear Stress Sensitizes Cancer Cells to Receptor-Mediated Apoptosis via Trimeric Death Receptors. *New J. Phys.* **15**, 015008.
- Nejman, D., Livyatan, I., Fuks, G., Gavert, N., Zwang, Y., Geller, L.T., Rotter-Maskowitz, A., Weiser, R., Malle, G., Gigi, E., et al. (2020). The human tumor microbiome is composed of tumor type-specific intracellular bacteria. *Science* **368**, 973–980.
- Nobes, C.D., and Hall, A. (1995). Rho, rac, and cdc42 GTPases regulate the assembly of multimolecular focal complexes associated with actin stress fibers, lamellipodia, and filopodia. *Cell* **81**, 53–62.
- Ocaña, O.H., Córcoles, R., Fabra, A., Moreno-Bueno, G., Acloque, H., Vega, S., Barrallo-Gimeno, A., Cano, A., and Nieto, M.A. (2012). Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer *Prrx1*. *Cancer Cell* **22**, 709–724.
- Ohgushi, M., Matsumura, M., Eiraku, M., Murakami, K., Aramaki, T., Nishiyama, A., Muguruma, K., Nakano, T., Suga, H., Ueno, M., et al. (2010). Molecular pathway and cell state responsible for dissociation-induced apoptosis in human pluripotent stem cells. *Cell Stem Cell* **7**, 225–239.
- Oksanen, J., Guillaume Blanchet, F., Friendly, M., Kindt, R., Legendre, P., McGlenn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, et al. (2019). *vegan: Community Ecology Package* (R package version 2.5-6). <https://CRAN.R-project.org/package=vegan>.
- Parhi, L., Alon-Maimon, T., Sol, A., Nejman, D., Shshadeh, A., Fainsod-Levi, T., Yajuk, O., Isaacson, B., Abed, J., Maalouf, N., et al. (2020). Breast cancer colonization by *Fusobacterium nucleatum* accelerates tumor growth and metastatic progression. *Nat. Commun.* **11**, 3259.
- Pertz, O., Hodgson, L., Klemke, R.L., and Hahn, K.M. (2006). Spatiotemporal dynamics of RhoA activity in migrating cells. *Nature* **440**, 1069–1072.
- Pinato, D.J., Howlett, S., Ottaviani, D., Urus, H., Patel, A., Mineo, T., Brock, C., Power, D., Hatcher, O., Falconer, A., et al. (2019). Association of Prior Antibiotic Treatment With Survival and Response to Immune Checkpoint Inhibitor Therapy in Patients With Cancer. *JAMA Oncol.* **5**, 1774–1778.
- Pushalkar, S., Hundeyin, M., Daley, D., Zambirinis, C.P., Kurz, E., Mishra, A., Mohan, N., Aykut, B., Usyk, M., Torres, L.E., et al. (2018). The Pancreatic Cancer Microbiome Promotes Oncogenesis by Induction of Innate and Adaptive Immune Suppression. *Cancer Discov.* **8**, 403–416.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F.O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596.
- Regmi, S., Fu, A., and Luo, K.Q. (2017). High Shear Stresses under Exercise Condition Destroy Circulating Tumor Cells in a Microfluidic System. *Sci. Rep.* **7**, 39975.
- 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 e712.
- Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140.
- Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**, e2584.
- Routy, B., Le Chatelier, E., Derosa, L., Duong, C.P.M., Alou, M.T., Daillère, R., Fluckiger, A., Messaoudene, M., Rauber, C., Roberti, M.P., et al. (2018). Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science* **359**, 91–97.
- Salter, S.J., Cox, M.J., Turek, E.M., Calus, S.T., Cookson, W.O., Moffatt, M.F., Turner, P., Parkhill, J., Loman, N.J., and Walker, A.W. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* **12**, 87.
- Scatena, C., Roncella, M., Di Paolo, A., Aretini, P., Menicagli, M., Fanelli, G., Marini, C., Mazzanti, C.M., Ghilli, M., Sotgia, F., et al. (2018). Doxycycline, an Inhibitor of Mitochondrial Biogenesis, Effectively Reduces Cancer Stem Cells (CSCs) in Early Breast Cancer Patients: A Clinical Pilot Study. *Front. Oncol.* **8**, 452.
- Sethi, V., Kurtom, S., Tarique, M., Lavania, S., Malchiodi, Z., Hellmund, L., Zhang, L., Sharma, U., Giri, B., Garg, B., et al. (2018). Gut Microbiota Promotes Tumor Growth in Mice by Modulating Immune Response. *Gastroenterology* **155**, 33–37 e36.
- Shaw, A.K., Halpern, A.L., Beeson, K., Tran, B., Venter, J.C., and Martiny, J.B. (2008). It's all relative: ranking the diversity of aquatic bacterial communities. *Environ. Microbiol.* **10**, 2200–2210.
- Shi, J., Wu, X., Surma, M., Vemula, S., Zhang, L., Yang, Y., Kapur, R., and Wei, L. (2013). Distinct roles for ROCK1 and ROCK2 in the regulation of cell detachment. *Cell Death Dis.* **4**, e483.
- Shi, Y., Zheng, W., Yang, K., Harris, K.G., Ni, K., Xue, L., Lin, W., Chang, E.B., Weichselbaum, R.R., and Fu, Y.X. (2020). Intratumoral accumulation of gut microbiota facilitates CD47-based immunotherapy via STING signaling. *J. Exp. Med.* **217**, e20192282.
- Sivan, A., Corrales, L., Hubert, N., Williams, J.B., Aquino-Michaels, K., Earley, Z.M., Benaymin, F.W., Lei, Y.M., Jabri, B., Alegre, M.L., et al. (2015). Commensal *Bifidobacterium* promotes antitumor immunity and facilitates anti-PD-L1 efficacy. *Science* **350**, 1084–1089.

- Smith, T., Heger, A., and Sudbery, I. (2017). UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.* *27*, 491–499.
- Spike, B.T., Engle, D.D., Lin, J.C., Cheung, S.K., La, J., and Wahl, G.M. (2012). A mammary stem cell population identified and characterized in late embryogenesis reveals similarities to human breast cancer. *Cell Stem Cell* *10*, 183–197.
- Strobel, M., Pfortner, H., Tuchscherer, L., Völker, U., Schmidt, F., Kramko, N., Schnittler, H.J., Fraunholz, M.J., Löffler, B., Peters, G., and Niemann, S. (2016). Post-invasion events after infection with *Staphylococcus aureus* are strongly dependent on both the host cell type and the infecting *S. aureus* strain. *Clin. Microbiol. Infect.* *22*, 799–809.
- Susaki, E.A., Tainaka, K., Perrin, D., Yukinaga, H., Kuno, A., and Ueda, H.R. (2015). Advanced CUBIC protocols for whole-brain and whole-body clearing and imaging. *Nat. Protoc.* *10*, 1709–1727.
- Suzuki, M.T., Taylor, L.T., and DeLong, E.F. (2000). Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5 $\phi$ -nuclease assays. *Appl. Environ. Microbiol.* *66*, 4605–4614.
- Thompson, K.J., Ingle, J.N., Tang, X., Chia, N., Jeraldo, P.R., Walther-Antonio, M.R., Kandimalla, K.K., Johnson, S., Yao, J.Z., Harrington, S.C., et al. (2017). A comprehensive analysis of breast cancer microbiota and host gene expression. *PLoS ONE* *12*, e0188873.
- Tsai, J.H., Donaher, J.L., Murphy, D.A., Chau, S., and Yang, J. (2012). Spatio-temporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis. *Cancer Cell* *22*, 725–736.
- Urbaniak, C., Cummins, J., Brackstone, M., Macklaim, J.M., Gloor, G.B., Baban, C.K., Scott, L., O'Hanlon, D.M., Burton, J.P., Francis, K.P., et al. (2014). Microbiota of human breast tissue. *Appl. Environ. Microbiol.* *80*, 3007–3014.
- Urbaniak, C., Gloor, G.B., Brackstone, M., Scott, L., Tangney, M., and Reid, G. (2016). The Microbiota of Breast Tissue and Its Association with Breast Cancer. *Appl. Environ. Microbiol.* *82*, 5039–5048.
- Vandeputte, D., Kathagen, G., D'hoë, K., Vieira-Silva, S., Valles-Colomer, M., Sabino, J., Wang, J., Tito, R.Y., De Commer, L., Darzi, Y., et al. (2017). Quantitative microbiome profiling links gut community variation to microbial load. *Nature* *551*, 507–511.
- Vétizou, M., Pitt, J.M., Daillère, R., Lepage, P., Waldschmitt, N., Flament, C., Rusakiewicz, S., Routy, B., Roberti, M.P., Duong, C.P., et al. (2015). Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. *Science* *350*, 1079–1084.
- Wang, W., Lin, L., Du, Y., Song, Y., Peng, X., Chen, X., and Yang, C.J. (2019). Assessing the viability of transplanted gut microbiota by sequential tagging with D-amino acid-based metabolic probes. *Nat. Commun.* *10*, 1317.
- Wickham, H. (2016). ggplot2.
- Xavier, J.B., Young, V.B., Skufca, J., Ginty, F., Testerman, T., Pearson, A.T., Macklin, P., Mitchell, A., Shmulevich, I., Xie, L., et al. (2020). The Cancer Microbiome: Distinguishing Direct and Indirect Effects Requires a Systemic View. *Trends Cancer* *6*, 192–204.
- Xuan, C., Shamonki, J.M., Chung, A., Dinome, M.L., Chung, M., Sieling, P.A., and Lee, D.J. (2014). Microbial dysbiosis is associated with human breast cancer. *PLoS ONE* *9*, e83744.
- Yu, T., Guo, F., Yu, Y., Sun, T., Ma, D., Han, J., Qian, Y., Kryczek, I., Sun, D., Nagarsheth, N., et al. (2017). *Fusobacterium nucleatum* Promotes Chemoresistance to Colorectal Cancer by Modulating Autophagy. *Cell* *170*, 548–563 e516.
- Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* *16*, 284–287. <https://doi.org/10.1089/omi.2011.0118>.
- Zackular, J.P., Baxter, N.T., Iverson, K.D., Sadler, W.D., Petrosino, J.F., Chen, G.Y., and Schloss, P.D. (2013). The gut microbiome modulates colon tumorigenesis. *MBio* *4*, e00692–e13.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
Mouse Anti-E. coli LPS	Abcam	Cat#ab35654; RRID:AB_732222
Anti-Gram Positive Bacteria antibody	Abcam	Cat#ab20344; RRID:AB_445526
Rat Anti-mouse F4/80 Clone BM8	Biolegend	Cat#123102; RRID:AB_893506
Pig Anti-Cytokeratin 8+18	Abcam	Cat#ab194130; RRID:AB_2728659
anti-mouse CD31	Biolegend	Cat#102501; RRID:AB_312908
Anti-Collagen I	Abcam	Cat#254113
Anti-RhoA	CST	Cat#8789
Anti-ROCK2	Abcam	Cat#ab228008
Anti-Rabbit IgG HRP-Linked antibody	Cell Signaling Technology	Cat#7074P2; RRID:AB_2099233
Anti-mouse IgG HRP-linked antibody	Cell Signaling Technology	Cat#7076P2; RRID:AB_330924
V450-anti-mouse CD45	BD	Cat#560501; RRID:AB_1645275
PerCP/Cy 5.5 antimouse CD3eClone 145-2c11	Biolegend	Cat#100328; RRID:AB_893318
APC anti-mouse CD49b	Biolegend	Cat#103511; RRID:AB_528830
PE/Cy7 anti-mouse NK-1.1 Antibody	Biolegend	Cat#108714; RRID:AB_389364
APC anti-mouse CD19 Antibody	Biolegend	Cat#115512; RRID:AB_313647
APC/Cyanine7 anti-mouse CD11c Antibody	Biolegend	Cat#117324 RRID:AB_830649
APC anti-mouse MHCII	Biolegend	Cat#107614; RRID:AB_313329
FITC anti-Ly6C	Biolegend	Cat#128006; RRID:AB_1186135
PE anti-mouse Ly-6G	Biolegend	Cat#127608; RRID:AB_1186099
PerCP/Cy 5.5 antimouse F4/80 clone BM8	Biolegend	Cat#123128; RRID:AB_893484
APC/Cy7 anti-mouse/human CD11b	Biolegend	Cat#101226; RRID:AB_830642
<b>Bacterial and virus strains</b>		
<i>S. xyloso</i>	This paper	N/A
<i>E. faecalis</i>	This paper	N/A
<i>L. animalis</i>	This paper	N/A
<i>S. cuniculi</i>	This paper	N/A
<i>S. sanguinis</i>	This paper	N/A
<b>Biological samples</b>		
Human breast tumor tissues	The First Affiliated Hospital of Zhejiang University	N/A
Human breast tumor tissues	Zhejiang Cancer Hospital	N/A
<b>Chemicals, peptides, and recombinant proteins</b>		
Vancomycin	Sangon Biotech	Cat#A600983-0001
Imipenem/cilastatin	Merck Sharp	N/A
Neomycin	Sangon Biotech	Cat#A610366-0025
Amphotericin	Sangon Biotech	Cat#A610030-0001
Doxycycline	Sangon Biotech	Cat#A600889-0100
Clarithromycin	Sangon Biotech	Cat#A504058-0025
Azithromycin	Sangon Biotech	Cat#A602222-0005
Ampicillin	Sangon Biotech	Cat#A100339-0025
Gentamicin	Sangon Biotech	Cat#A620217-0005
2X SSC buffer	Ambion	Cat#AM9765
DMEM/F12 media	Gibco	Cat#C11330500CP

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Collagenase type 3	Worthington Biochemical	Cat#LS004182
Hyaluronidase	Worthington Biochemical	Cat#LS002592
ACK	Beyotime biotechnology	Cat#C3702
FBS	Jackson	Cat#017-000-121
Trypsin-EDTA	Thermo	Cat#25200072
DNaseI	Worthington Biochemical	CatLS002139
Columbia blood agar	OXOID	Cat#CM0331B
Sheep blood	Solarbio	Cat#TX0030
BHI (brain heart infusion)	Solarbio	Cat#LA0360
Schaedler anaerobe agar	OXOID	Cat#CM0437
Lysozyme	Sigma	Cat#62970
Lysostaphin	Sigma	Cat#L7386
Proteinase K	Ambion	Cat#AM2546
Formamide	Ambion	Cat#AM9342
Dextran sulfate	Sigma	Cat#D8906
<i>E. coli</i> tRNA	Sigma	Cat#R4251
BSA	Ambio	Cat#AM2616
DAPI	Sigma	Cat#D9564
Antifade Mounting Medium	Beyotime	Cat#P0126
Alexa Fluor 488-phalloidin	Invitrogen	Cat#A12379
CellTrace™ Far Red Cell Proliferation staining	Invitrogen	Cat#C34564
Lysis/Binding/Wash Buffer plus Aprotinin	Sigma	Cat#A1153-25MG
Pepstatin A	Sigma	Cat#P8-25MG
EGF	Thermo	Cat#PMG8043
R-spondin1	R&D	Cat#7150-RS-250/CF
ROCK inhibitor Y27632	TOCRIS	Cat#1254/10
CFSE	Invitrogen	Cat#65-0850-84

**Critical commercial assays**

Histostain-Plus IHC Kit	NEOBIOSCIENCE	Cat#ENS004.120
LIVE/DEAD ®BacLight™ Bacterial Viability and Counting Kit	Invitrogen	Cat#L34856
Active Rho Detection Kit	CST	Cat#8820S
QIAamp DNA Microbiome kit	QIAGEN	Cat# 51704
QIAamp PowerFecal (pro) DNA kit	QIAGEN	Cat#51804
Takara2 Premix Ex Taq	Takara	Cat#RR390A
TruePrep® Index Kit V3 for Illumina®	Vazyme	Cat#TD203
KAPA Hyper Prep Kits	KAPA	KK8505

**Deposited data**

16S amplicon sequencing	BioProject: PRJNA681060	<a href="https://dataview.ncbi.nlm.nih.gov/object/PRJNA681060?reviewer=7v2h87ups0iauqrpdmae8ue64j">https://dataview.ncbi.nlm.nih.gov/object/PRJNA681060?reviewer=7v2h87ups0iauqrpdmae8ue64j</a>
Single cell RNAseq	BioProject: PRJNA681289	<a href="https://dataview.ncbi.nlm.nih.gov/object/PRJNA681289?reviewer=hkc97qrfssn7b4o3jfr22o0i8p">https://dataview.ncbi.nlm.nih.gov/object/PRJNA681289?reviewer=hkc97qrfssn7b4o3jfr22o0i8p</a>

**Experimental models: Organisms/strains**

FVB/N-Tg(MMTV-PyVT)634Mul/J	The Jackson Laboratory	Cat#002374
Fvb	Shanghai SLAC Laboratory Animal Company	N/A

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
NPSG (NOD-Prkdc <sup>scid</sup> Il2rg <sup>null</sup> /Shjh)	Shanghai Jihui Laboratory Animal Care Co.,Ltd	N/A
mTmG mice (B6.129(Cg)-Gt(ROSA)26Sor <sup>tm4(ACTB-tdTomato,-EGFP)Luo/J</sup> )	The Jackson Laboratory	Cat#007676
MMTV-Wnt transgenic mice (FVB.Cg-Tg(Wnt1)1Hev/J)	The Jackson Laboratory	Cat#002934
Germ-free FVB/N mice	GemParmatech	N/A
<b>Oligonucleotides</b>		
List of oligonucleotides	This paper, <a href="#">Table S6</a>	N/A
<b>Software and algorithms</b>		
GraphPad Prism, version 8	SCR_002798	<a href="https://www.graphpad.com/scientific-software/prism/">https://www.graphpad.com/scientific-software/prism/</a>
FlowJo, version 10	SCR_008520	<a href="https://www.flowjo.com/">https://www.flowjo.com/</a>
SnapGene v2.3.2	SCR_015052	<a href="https://www.snapgene.com/">https://www.snapgene.com/</a>
ImageJ	SCR_003070	<a href="https://imagej.nih.gov/ij/">https://imagej.nih.gov/ij/</a>
TrimGalore	Babraham Institute	<a href="https://github.com/FelixKrueger/TrimGalore">https://github.com/FelixKrueger/TrimGalore</a>
FastQC	Babraham Institute	<a href="https://www.bioinformatics.babraham.ac.uk/projects/fastqc/">https://www.bioinformatics.babraham.ac.uk/projects/fastqc/</a>
Seurat v4	Satija Lab	<a href="https://github.com/satijalab/seurat">https://github.com/satijalab/seurat</a>
clusterProfiler	<a href="#">Yu et al., 2012</a>	<a href="https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html">https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html</a>
UMI-tools	<a href="#">Smith et al., 2017</a>	<a href="https://github.com/CGATOxford/UMI-tools">https://github.com/CGATOxford/UMI-tools</a>
Imaris 9.3.1	Bitplane	<a href="https://imaris.oxinst.com/versions/9-3">https://imaris.oxinst.com/versions/9-3</a>
vsearch v2.14.2	<a href="#">Rognes et al., 2016</a>	N/A
usearch v10	<a href="#">Edgar, 2010</a>	N/A

**RESOURCE AVAILABILITY**

**Lead contact**

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Shang Cai ([caishang@westlake.edu.cn](mailto:caishang@westlake.edu.cn)).

**Materials availability**

All requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact ([caishang@westlake.edu.cn](mailto:caishang@westlake.edu.cn)). All reagents including antibodies, bacteria and plasmid may be available on request after completion of a Materials Transfer Agreement.

**Data and code availability**

16S amplicon sequencing and Single-cell RNA-seq data have been deposited at SRA and are publicly available as of the date of publication. Project number and accession links are listed in the [key resources table](#).

This paper does not report original code.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

**EXPERIMENTAL MODEL AND SUBJECT DETAILS**

**Mice**

Female MMTV-PyMT transgenic mice (FVB/N-Tg(MMTV-PyVT)634Mul/J, 002374), which develop spontaneous breast tumor, were purchased from The Jackson Laboratory (The Jackson Laboratory, Bar Harbor, Maine, USA) and bred in the laboratory animal resources center of Westlake University. Female Fvb mice, 6–8 weeks old were purchased from Shanghai SLAC Laboratory Animal Company (Shanghai, China). NPSG (NOD-Prkdc<sup>scid</sup> Il2rg<sup>null</sup>/Shjh) mice were purchased from Shanghai Jihui Laboratory Animal Care Co.,Ltd. mTmG mice (B6.129(Cg)-Gt(ROSA)26Sor<sup>tm4(ACTB-tdTomato,-EGFP)Luo/J</sup>, 007676) were purchased from The Jackson Laboratory (The Jackson Laboratory, Bar Harbor, Maine, USA). Female germ-free FVB/N mice (6–8 weeks) were bred and maintained in

special plastic isolators (GemParnatech, Nanjing, China). MMTV-Wnt tumor cell were digested from MMTV-Wnt transgenic mice (FVB.Cg-Tg(Wnt1)1Hev/J, 002934), which were purchased from the Jackson Laboratory (The Jackson Laboratory, Bar Harbor, Maine, USA).

Animals were housed in a specific pathogen-free conditions and fed standard mouse chow. All animal experiments were carried out in compliance with China laws and regulations. The local institutional animal ethics board approved all mouse experiments (permission numbers: 19-001-CS). Experiments were performed in accordance with government and institutional guidelines and regulations.

### Human samples collection

Human tissues were collected in the sterile surgery room from two medical centers from The First Affiliated Hospital of Zhejiang University and Zhejiang Cancer Hospital respectively. Fresh tissues of normal breast, breast tumor and lymphoid metastasis were immediately transferred to germ free 15ml conical tubes with sterile DMEM culture medium. Samples were processed in the clean and sterile cell culture hood with autoclaved dissection tools. All samples were collected and analyzed after informed consent was obtained from the patients and according to IRB-approved protocols: IRB-2019-99 and IRB-2020-634. The age and gender information of all human samples was provided in Table S7. All of these human samples were extracted DNA for bacterial quantification and 16S library preparation. The samples of bacteria biomass lower than EBC were excluded. Therefore, 6 of normal breast tissues, 11 of breast tumors and 4 of lymphoid metastases were analyzed.

## METHOD DETAILS

### Mouse mammary tumor dissection and dissociation

Mouse breast tumors were dissected and digested according to our previous protocol with minor revisions (Cai et al., 2017). Mouse breast tumors were dissected into 9.5mL DMEM/F12 media (C11330500CP, Gibco), the tumor tissues were chopped with razor blade into 1mm small pieces, followed by digestion with collagenase type 3 (300U/mL) (Worthington Biochemical #LS004182) and hyaluronidase (100U/mL) (Worthington Biochemical #LS002592) at 37°C for 2 h. The digested tissues were pooled into a 50mL conical and spun down at 1500rpm for 5min. Red blood cells were lysed with 5 mL ACK (Beyotime biotechnology #C3702-120mL) on ice for 5 min, washed once with HBSS+2%FBS (Jackson #017-000-121) and then digested with 5 mL pre-warmed 0.25% trypsin-EDTA (Thermo #25200072) at 37°C for 5 min by mechanical pipetting for dissociation. Digested cells were washed once with HBSS+2%FBS (Jackson #017-000-121), then digested with 0.1 mg/mL DNaseI (Worthington Biochemical LS002139) in DMEM/F12 media at 37°C for 5 min, washed once and filtered through a 40 $\mu$ m mesh. Dissociated cells were counted and resuspended in designated medium for subsequent staining and FACS analysis.

### Bacteria culture and identification

For isolation of anaerobic or aerobic bacteria, tissue pieces (around 0.25g) were homogenized with glass homogenizer in 1 mL ice-cold PBS under sterile conditions. PBS was used as tissue surrogate and went through the same workflow to evaluate the environmental contaminants. For aerobic culture, 100 $\mu$ L sample homogenate was plated on Columbia blood agar (OXOID-#CM0331B) +5% sheep blood (Solarbio-#TX0030), Man Rogosa Sharpe Medium (M0303) and BHI (brain heart infusion) (Solarbio-#LA0360) at 37°C aerobically with 5% CO<sub>2</sub>. For anaerobic culture, 100 $\mu$ L sample homogenate was plated on Schaedler anaerobe agar (OXOID-#CM0437) in an anaerobic chamber Hypoxystation (SIA-CC002). The plates were incubated at 37°C for either 3 days in aerobic conditions or for 5 days in anaerobic conditions.

For identification of bacteria strain, colonies were picked and streaked in designated plate and condition for 1-3 days to get single colonies. The single colony was picked to grown in liquid medium and run colony PCR subsequently. Briefly, the 15  $\mu$ L reaction mix contained 2X Es Taq MasterMix (Dye) (CW bio-#CW0690), 200 nM primers (27F: 5'-AGRGTGGATCMTGGCTCAG -3'; 1492R: 5'-TACGGYTACCTTGTTAYGACTT-3') (Shaw et al., 2008) (Table S6). The reaction was programmed as follows: 95°C for 10 minutes, 10 cycles of 95°C for 15 s, reduced annealing temperature for 15 s (63.4 °C, 1°C /every second cycle) and 72°C for 60s. Then 30 cycles of 95°C for 15 s, 55°C for 15s, 72°C for 60s. Final extension reactions were carried out for 10 min at 72°C. The PCR product was sent out for sequencing and sequencing results were aligned to the 16S rRNA sequences (Bacteria and Archaea) database in the NCBI blastn site.

### Quantification and profiling of tissue-resident microbiome

Different from the gut microbiome, which contains about 10<sup>10</sup> bacteria/gram, tissue-resident microbiota has a much lower bacteria load by several orders of magnitude and a seriously higher host genome contamination, therefore is more challenging to quantify the absolute abundance and to profile the microbial community. In addition, when bacteria load drops down to a certain level, the environmental contamination becomes a major issue that can mask the real microbiota signal (Eisenhofer et al., 2019). To overcome these problems, and to accurately and sensitively detect and profile tissue resident microbiota, we carefully optimized the whole experimental procedures from beginning to the end as listed below. With this optimized method, we were able to faithfully quantify and construct the 16S library for as low as 10<sup>4</sup> equivalent bacteria/gram tissue, which is roughly 1 bacteria out of 10,000 cells.

### Mouse sample collection and human sample collection

Tumor bearing mice were housed in the SPF level Westlake Animal Facility. Tumor dissection and processing were strictly carried out in the clean and sterile cell culture hood with autoclaved dissection tools. We always use fresh tissue for microbiota analysis to avoid fixation caused quality drop and artifacts.

Human samples were collected in the sterile surgery room. Fresh tissues were immediately transferred to germ free 15ml conical tubes with sterile DMEM culture medium. Samples were processed in the clean and sterile cell culture hood with autoclaved dissection tools. We always use fresh tissue for microbiota analysis to avoid fixation caused quality drop and artifacts.

### Bacteria DNA extraction

There are three major issues affecting the bacteria DNA extraction: enormous host genome, bacteria DNA release and contaminations in the reagents. We tested different DNA extraction kit, proper tissue amount, beads shearing to get high quality bacteria DNA. The contamination in the reagent is hard to eliminate, therefore, we set up stringent environmental controls that underwent exact the same procedure with the samples to generate a contamination landscape.

Whether to eliminate host cell genome: QIAamp DNA Microbiome kit (QIAGEN # 51704), which can deplete host DNA, and QIAamp PowerFecal (pro) DNA kit (QIAGEN #51804) were tested for DNA extraction. All the tested kits contained bead-beating step and the bead-beating steps were performed in a TissueLyser II (QIAGEN, 85300) for 10 min at 30 Hz. The bacterial DNA quantity decreased without the bead-beating step. For the QIAamp DNA Microbiome kit, it contained steps to lysis host cells with buffer AHL and Benzonase to digest host DNA, the extra steps to lyse the host genome always significantly decrease the sample yield for tissues, which eventually lower the sensitivity of the qPCR quantification, and frequently generate ambiguous results. Therefore, we determine to use QIAamp PowerFecal (pro) DNA kit with no host genome elimination step to extract bacteria DNA for subsequent analysis.

Tissue processing: 200mg tumor tissue was determined to be used to extract DNA, more tissue would decrease the efficiency of extraction due to limited binding capacity of spin column. Tumor tissue samples were grinded in 1ml sterile PBS. Centrifuge at full speed for 10 minutes and discard the supernatant. The pellet could be extracted or frozen at  $-80^{\circ}\text{C}$ .

To investigate the tumor cell associated microbiota, tumor tissues were digested according to the tumor digestion session in Materials and Methods and around  $5 \times 10^7$  cells were used to extract total DNA. Total genomic DNA was extracted from tumor tissue, tumor cells and fecal samples with the QIAamp PowerFecal (pro) DNA kit (QIAGEN-#51804) according to the manufacturer's instructions. Samples were disrupted using PowerBead Pro Tubes included in the kit in a TissueLyser II for 10 min at 30 Hz. The PowerBead Pro Tubes contain large beads. Then the DNA was purified following manufacturer's instruction.

Set up controls: As contaminations in the reagents are not easy to eliminate, it is essential to set up proper negative controls. We use PBS as a tissue surrogate to undergo the same processing steps as tissue samples. The final DNA product was used as the environmental bacteria control (EBC). At the same time, we use pure DEPC treated water as the NTC when running the qPCR quantification assay.

### qPCR quantification

As qPCR reagents also contain certain levels of bacteria DNA contaminations, we set out to screen the best commercialized qPCR kit with best sensitivity (the lowest amount of bacteria DNA it can faithfully detect), best specificity (Taqman probe method has higher specificity than SYBRgreen), and best stability (smallest variation between various experiments). (Figure S1 A-C)

Several qPCR reaction mixtures, both SYBRgreen and Taqman probe mixture, were tested for bacterial quantification. Takara2, TOYOBO2, Toroid2 were Taqman probe mixture and 2X ChamQ Vazyme, YEASEN Hieff, YEASEN UNICON were SYBR green mixture. V9 region of the 16S ribosomal RNA were amplified with following primers: Forward 5'-CGGTGAATACGTTTCYCGG-3', Reverse 5'-GGWTACCTTGTACGACTT-3', and Probe 5'-CTTGTACACACCGCCCGTC-3' (Suzuki et al., 2000) (Table S6). *Escherichia coli* DNA was used to plot a standard curve to calculate bacterial DNA concentration in the sample and NTCs were included for the reactions (Ibekwe et al., 2002). The comparison between various commercialized qPCR kit showed that Takara2 Premix Ex Taq (Takara-#RR390A) kit has the lowest bacteria DNA contamination, highest sensitivity, specificity and stability. (Figure S1 A-C)

For qPCR quantification, briefly, 10  $\mu\text{L}$  reaction mix containing Premix Ex Taq (probe qPCR), 750 nM of forward primer, 500 nM reverse primer and 250nM probe, and 1  $\mu\text{L}$  sample DNA was loaded on the qTOWEP384/G-Analytik Jena real-time system. The reaction was programmed as follows: denaturation at  $94^{\circ}\text{C}$  for 10 minutes, 40 cycles of  $94^{\circ}\text{C}$  for 15 s,  $60^{\circ}\text{C}$  for 60 s. Raw threshold cycle (Ct) values were normalized according to a bacterial standard curve produced with *E.coli* DNA.

### 16S library construction

When the bacteria biomass drops down to certain level, the 16S library mainly contains bacteria signals coming from the reagent contamination. To faithfully and accurately construct the 16S library, representing the sample microbiota profile, we optimized the library construction methods as exemplified in Figure S1 F-H. The V4 hypervariable region of the 16S gene was amplified from the genomic DNAs of mice fecal samples and tumor samples according to the Illumina 16S metagenomics protocol (Part #15042322).

For gut microbiome library construction, V4 region was initially amplified using the primer set 515F (0.4 $\mu\text{M}$ , 5 $\phi$ -TCGTCCGGCAGCGTCAGATGTGTATAAGAGACAGGTGYCAGCMGCCGCGGTAA -3 $\phi$ ) and 805R (0.4 $\mu\text{M}$ , 5 $\phi$ -GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGGACTACNVGGGTWTCTAAT-3 $\phi$ ) (Table S6), each with overhang adapter sequences (IDT) using 2 $\times$  Kapa HiFi Hotstart ReadyMix DNA polymerase (KapaBiosystems). Amplification was performed with an initial heating step of  $95^{\circ}\text{C}$  for 3 minutes, 25 cycles of 30 seconds at  $95^{\circ}\text{C}$ , 30 seconds at  $55^{\circ}\text{C}$ , 30 seconds at  $72^{\circ}\text{C}$ , and final extension of  $72^{\circ}\text{C}$  for 5 minutes. After amplification, for fecal sample, AMPure XP beads (Beckman Coulter, #A63881) were used to purify the PCR products following manufacturer's instructions. Purified PCR products were subjected to second round index PCR. Specifically, dual indices from

TruePrep® Index Kit V3 for Illumina® (5ul for each index, Vazyme #TD203) were added to target amplicons in a second PCR using 2× Kapa HiFi Hotstart ReadyMix DNA polymerase (25ul). PCR condition was 3 minutes at 95°C, with 8 cycles of 30 seconds at 95°C, 30 seconds at 55°C, 30 seconds at 72°C, and final extension of 5 minutes at 72°C. Libraries were purified by the AMPure XP (0.8X) bead before quantification by using the Qubit DNA assay (Thermo Fisher Scientific) and qPCR, or quality check by Fragment Analyzer-12/96 (GENE-QC006).

For tumor-resident microbiota library construction, the V4 hypervariable region of the 16S gene was amplified from tissue samples according to the previous method with some optimizations (Riquelme et al., 2019). We used PBS that underwent all the sample preparation procedures as environmental negative control and DEPC-treated water as no template control, as well as *E. coli* as the positive control. The V4 hypervariable region of 16S was amplified with biotinylated primer set 515F (0.4uM, 5c- TCGTCGGCAGCGTC AGATGTGTATAAGAGACAGGTGYCAGCMGCCGCGGTAA -3c) and 805R (0.4uM, 5c- GTCTCGTGGGCTCGGAGATGTGTATAA GAGACAGGGACTACNVGGGTWTCTAAT-3c) (Table S6), each with overhang adapter sequences (IDT), using 2× Kapa HiFi Hotstart ReadyMix DNA polymerase (KapaBiosystems). Amplification was performed with an initial heating step of 95°C for 3 minutes, 25 cycles of 30 seconds at 95°C, 30 seconds at 55°C, 30 seconds at 72°C, and final extension of 72°C for 5 minutes. The PCR cycles were optimized to 30 cycles. PCR products were then purified and enriched by Dynabeads™ MyOne™ Streptavidin C1 Beads (Thermo Fisher Scientific, #5002) following manufacturer's instructions. Biotin beads containing first round PCR products were directly subjected to the second round index PCR. Dual indices from TruePrep® Index Kit V3 for Illumina® (5ul for each index, Vazyme #TD203) were added to target amplicons in a second PCR using 2× Kapa HiFi Hotstart ReadyMix DNA polymerase (25ul). PCR conditions were 3 minutes at 95°C, with 8 cycles of 30 seconds at 95°C, 30 seconds at 55°C, 30 seconds at 72°C, and final extension of 5 minutes at 72°C. Libraries were purified by the AMPure XP (0.8X) bead before quantification by using the Qubit DNA assay (Thermo Fisher Scientific) and qPCR, or quality check by Fragment Analyzer-12/96 (GENE-QC006). Negative controls including no template controls and environmental controls were included in all sequencing runs. Samples were sequenced on the Novaseq with 2X250 bp paired-end reads (Novogene, Tianjin).

## Microbiome data analysis

### Processing metagenomic sequencing data

The sequencing data were split using de-multiplexing tools bcl2fastq2 (double check with sequencing service company for de-multiplexing method they used) to generate fastq format files for every sample using barcode sequences. All reads were analyzed using a standardized metagenome bioinformatics pipeline using vsearch v2.14.2 (Rognes et al., 2016) and usearch v10 (Edgar, 2010). First, paired-end reads were merged using "vsearch -fastq\_mergepairs" with default parameters. Then, adaptor trimming and low quality read removal were carried out by "vsearch -fastq\_filter" with maximum expected error rate 0.01. After quality control filtering, a total of 67170279 reads were processed. On average, fecal samples had 1405670 reads per sample, mouse tissue samples had 230468 reads per sample and human samples had 174804 reads per sample. Metagenome template were constructed using the quality control passed reads. Reads were dereplicated using "vsearch -derep\_fulllength" with singleton reads discarded. All remaining unique sequences were denoised by unoise3 method to get candidate sequence features, which balances between resolution and specificity (Quast et al., 2013). Lastly, chimeric features were removed using "vsearch -uchime\_ref" against the SILVA reference release123 (Quast et al., 2013).

The abundance table were generated by mapping all quality control passing reads onto the template sequence features using "vsearch -usearch\_global" with identity threshold 0.97. Sequence features were classified by "vsearch -sintax" (Edgar, 2016) with cutoff set to 0.6. Non-specific sequence features from eukaryote, chloroplast and mitochondria were filtered out. Only samples with enough prokaryote specific reads were kept using the cutoff of 3000 and 1000 for mouse and human samples respectively. All kept samples were normalized to the same number of reads for a fair comparison (3000 reads for mouse and 1000 reads for human) using the "rrarefy" function of R package "vegan" (Oksanen et al., 2019).

### Contamination correction

The contamination effect in each mouse sample was numerically corrected using the data from qPCR quantification experiment. First, the relative abundance of each bacterial species was obtained from the metagenomic sequencing data. Then, absolute bacterial amount per species per sample was calculated according to the qPCR quantification value and relative abundance within the sample. The contamination bacteria from background environment was measured as the median absolute amount of a species among all negative control samples, formulated as below:

$$C_{ij} = c_{ij} * Q_i$$

$$C_j = \text{Median}(C_{ij})$$

where  $c_{ij}$  is the relative abundance of species  $j$  from sample  $i$ , i.e. the percentage of species  $j$  data among the data of all species in sample  $i$ ;  $Q_i$  is the qPCR quantification result of sample  $i$  in unit of CFU/g.  $C_j$  is the overall contamination effect of species  $j$ , as the median relative abundance of species  $j$  among all samples.

The correction simply took off the overall contamination effect of species  $j$  from the measured species  $j$  in each sample:

$$A_{ij} = a_{ij} * Q_i - C_j$$

where  $a_{ij}$  is the relative abundance of species  $j$  from sample  $i$ ;  $Q_i$  is quantification of total bacteria of sample  $i$ ;  $A_{ij}$  is the corrected amount for species  $j$  of sample  $i$ .

For human samples, contamination correction was based on binomial tests of every taxon's abundance between samples and negative controls. The probability  $p$  of binomial distribution was estimated as the taxon's occurrence frequency in negative control samples. The number of occurrence samples and the total number of samples were assigned as  $x$  and  $n$  in the binomial test respectively. Taxon with  $p$ -value < 0.05 were kept in the following analysis.

#### **Bioinformatic analysis for metagenomics data**

Contamination corrected abundance data were input into edgeR (Robinson et al., 2010) for differential analysis between two or more sample of origins. Taxonomies with FDR < 0.25 were considered as enriched or depleted. Volcano plots were taken from edgeR output. Heatmaps were generated by R package "pheatmap" (Kolde, 2019) on the contamination corrected abundance using Manhattan distance. The mean and standard error of abundance of each cluster were presented as bar plot

Alpha diversity matrixes and correspondent rarefaction analysis were calculated by "-alpha\_div" and "-alpha\_div\_rare" in usearch, using contamination corrected abundance. Wilcoxon test was used to evaluate the significance of alpha diversity difference between two groups.

Beta diversity matrixes were calculated by "-beta\_div" in usearch. Principal Coordinate Analysis (PCoA) and Constrained Principal Coordinate Analysis (CPCoA) were performed by function "beta\_pcoa" and "beta\_cpcoa\_dis" in R package "amplicon" (Liu et al., 2021) using default parameter. To compare the differences in the community composition between groups, Multiple Response Permutation Procedure (MRPP) test was performed by R package "vegan".

The stack plot of taxonomy was generated by "tax\_stackplot" function of R package "amplicon" based on relative abundance with the top taxonomies selected by the mean abundance of all samples in the analysis.

For microbiome phenotypes prediction, abundance table were regenerated using "vsearch -usearch\_global" (identity cutoff 0.97) and reference database Greengene release 13\_8 (DeSantis et al., 2006). The abovementioned contamination correction method was applied to the abundance table. Bugbase (Ward et al. 2017) was used for phenotype prediction. The Wilcoxon test was applied to evaluate the significance of differences between groups on given phenotypes.

To evaluate the taxonomy difference among sample groups, a straightforward group mean method on abundance tables was used to generate one abundance table for each sample group. Beta diversity analysis was applied to the group abundance table for the distance matrix between groups. Then, the distance matrix was transformed into the similarity matrix using the following reciprocal transformation:

$$S_{ij} = 1 / (1 + D_{ij})$$

where  $S_{ij}$  was the similarity between the two groups and  $D_{ij}$  was the beta diversity distance. The similarity matrix was used as input to R package "pheatmap" for heatmap and clustering analysis.

#### **Antibiotics treatments on MMTV-PyMT tumor cells**

Dissociated spontaneous PyMT tumor cells were cultured in the presence of Ampicillin+Gentamicin (200 µg/ml+200 µg/ml) or Doxycycline (10 µg/ml) overnight at 37 °C and were subsequently lysed and spread onto the CBA agar plate. Ampicillin+Gentamicin were used to eliminate extracellular bacteria and Doxycycline was used to eliminate both extracellular and intracellular bacteria.

#### **Antibiotics treatments on MMTV-PyMT mouse model**

To ablate both the gut microbiome and the tumor microbiome, 8-week-old mice were administered an antibiotic cocktail (ATBx) as described (Iida et al., 2013) with some modifications. Mice started to be administered with high dose ATBx (500 µL/mouse), containing vancomycin (50 mg/mL; Sangon Biotech), imipenem/cilastatin (25 mg/mL; Merck Sharp&Dohme Corp.U.S.A.), neomycin (10 mg/mL; Sangon Biotech), and amphotericin (1 mg/mL; Sangon Biotech), by oral gavage daily for five consecutive days. Then the mice were treated with a low dose antibiotic solution (ATBx) containing vancomycin (0.5 mg/mL; Sangon Biotech), imipenem/cilastatin (0.5 mg/mL; Merck Sharp&Dohme Corp.U.S.A.), neomycin (1 mg/mL; Sangon Biotech), and amphotericin (0.5 µg/mL; Sangon Biotech), by the sterile drinking water until end points. Solutions and bottles were changed every other day due to the short half-life of imipenem.

Antibiotic activity was confirmed weekly by cultivating fecal pellets resuspended in PBS on CBA (Columbia Blood Agar Base (OXOID #CM0331B) with 5% sterile defibrinated sheep blood (Solarbio-#TX0030)) plates or SAA (OXOID-#CM0437) for 48h at 37°C in aerobic and anaerobic conditions.

To specifically eliminate intratumor microbiota, we selected tetracycline (Doxycycline 0.2 mg/mL; Sangon Biotech) and two other macrolide antibiotics (Clarithromycin 0.2 mg/mL; Sangon Biotech and Azithromycin 0.2 mg/mL; Sangon Biotech) and administered into mice through drinking water. These antibiotics are all spectral antibiotics and are known to be active against obligate intracellular and facultative intracellular organisms. In an alternative method, mice were injected intravenously with 300 µL ATBx antibiotics suspension (vancomycin (10 mg/mL; Sangon Biotech), imipenem/cilastatin (4 mg/mL; Merck Sharp&Dohme Corp.U.S.A), neomycin (1.5 mg/mL; Sangon Biotech)) every 48h until 48h before the final analysis end point.

To enforce the entry of bacteria into tumor cells by AG-IV administration, we treated MMTV-PyMT mTmG mice bearing PyMT-mTmG tumors with 200ul PBS (as control) or Ampicillin (5mg/ml) +Gentamicin (0.5mg/ml) via tail vein injection when tumors are palpable. Antibiotics were administered once every 3days for approximately 40 days. At the experiment end point, mice were sacrificed for bacteria quantification, tumor weight quantification and metastasis quantification.

### Antibiotics Treatment on Germ Free Mouse

Germ-free FVB/N mice (6–8 weeks) were bred and maintained in special plastic isolators (GemParmatech, Nanjing, China) and housed under a strict 12:12 hour light-dark cycle (lights on at 08:00). Animals were supplied with a 50-kGy irradiated sterile pelleted normal chow diet (Xietong Shengwu, Nanjing, China) and autoclaved tap water ad libitum. Bedding was replaced in all experiments every 7 days. All germ-free mice were tested weekly for fecal bacteria, viral, and fungus contamination by facility staff to ensure that GF unit was indeed sterile.

We first screened a breast tumor from spontaneous tumor model MMTV-PyMT-mTmG mice that contains regular level of tumor-resident microbiota. The presence of bacteria in the tumor was confirmed by culturing. Then we grafted the tumor cells carrying bacteria orthotopically onto germ free mice at  $2 \times 10^6$  cells/mouse. Antibiotic (ATBx in drinking water containing vancomycin (0.5 mg/mL; Sangon Biotech), imipenem/cilastatin (0.5 mg/mL; Merck Sharp&Dohme Corp.U.S.A.), neomycin (1 mg/mL; Sangon Biotech), and amphotericin (0.5  $\mu$ g/mL; Sangon Biotech)) or normal drinking water were subsequently supplied to the mice. Tumor growth was monitored twice a week and fecal bacteria were monitored every week until the end of the experiment. Germ-free mice were shipped back to the lab overnight at the end point of experiment and were euthanized for primary tumor and lung metastasis quantification.

### Injection of intratumor bacteria to MMTV-PyMT and MMTV-Wnt mTmG tumor

For MMTV-PyMT spontaneous mouse model, intratumor microbiota were first eliminated by Doxycycline 1 mg/ml treatment for 5 days when tumors reach 5mm in diameter. Recipient mice were recovered for one day for turnover of Doxycycline. Then  $2 \times 10^6$  certain strains of bacteria were directly injected into the PyMT spontaneous tumors once every 10 days for three times. 30 days later, mice bearing PyMT tumors were sacrificed to analyze lung metastasis. Bacteria used were *E. faecalis*, *S. cuniculi*, *S. xylosum* and *L. animalis*. For MMTV-Wnt tumor, intratumor microbiota were first eliminated by Clarithromycin 0.2 mg/ml treatment for 5 days when tumors reached 5mm in diameter. Recipient mice were recovered for another 3 days for turnover of Clarithromycin. Then  $2 \times 10^6$  certain strains of bacteria were directly injected into the Wnt primary tumors for only one shot. 3 weeks later, mice bearing Wnt tumors were sacrificed to analyze lung metastasis. Bacteria used were *E. faecalis*, *S. cuniculi*, *S. sanguinis*, *S. danieliae*, *S. xylosum* and *L. animalis*.

### 16S RNA FISH Assay

The EUB338 16S rRNA gene probe (GCTGCCTCCCGTAGGAGT) (Table S6) labeled with the fluorophore Cy5 were used to detect the bacterial colonization within human and mouse tissues by FISH. Nonspecific complement probe (CGACGGAGGGGCATCCTCA) (Table S6) was used as a control for hybridization protocol. The protocol was adapted from Leore T. Geller et al (Geller et al., 2017) and optimized with two additional steps to fully lyse Staphylococcus. FFPE blocks were sectioned into 5- $\mu$ m slices and deparaffinized by immersing slides in 100% xylene for 10 min, fresh 100% xylene for 5 min, and 100% ethanol, fresh 100% ethanol and 95% ethanol, each for 10min, finally in 70% ethanol and kept at 4°C for at least 2 hours. They were subsequently incubated in 2X SSC buffer (Ambion #AM9765) for 10min at RT, then samples were incubated in 1mg/mL lysozyme (62970, Sigma) at RT for 10 min, followed by treatment of 0.05 mg/mL lysostaphin (L7386, Sigma) at RT for 10min. Then treated with 10  $\mu$ g/mL proteinase K (Ambion #AM2546) for 10 min at RT. Samples were incubated twice with 2X SSC buffer for 5 min at RT, rinsed with a wash buffer containing 25% formamide (Ambion #AM9342) in 2X SSC buffer, then incubated with fresh wash buffer for 5 min at RT. Probes were diluted to 1ng/ $\mu$ L in hybridization buffer containing 25% formamide, 10% dextran sulfate (Sigma #D8906), 1mg/mL *E.coli* tRNA (Sigma #R4251), 0.02% BSA (Ambio #AM2616) and 2X SSC buffer. The probes were hybridized to the tissue overnight at 37°C. Unbound probes were washed off by wash buffer for 30 min at 37°C. Samples were then stained with 1 ng/mL DAPI (Sigma #D9564) for 5min at RT. Samples were mounted with Antifade Mounting Medium (Beyotime #P0126). Images were acquired on an inverted epifluorescence microscope (Nikon Eclipse Ti2).

### Immunofluorescence assays

For frozen section, tissue was fixed in 4% paraformaldehyde (BBI Life Sciences #E672002-0500) for 2hrs at 4°C, then immersed in 30% sucrose for infiltration overnight. Embed the fixed tissue in O.C.T. (Tissue-Tek) compound and freeze at -80 degree. For FFPE, tissue was fixed in 4% paraformaldehyde overnight at 4°C. Antigens were retrieved in slight boiling citrate buffer (10mM Sodium Citrate, 0.05% Tween 20, pH6.0) for 15min in microwave with low-to-medium power. Samples were blocked with TBS+2% BSA+5% Donkey serum+0.1% Triton X-100 for 1hr at room temperature in humid chamber. Subsequently, the samples were stained with primary antibodies: Mouse Anti-E. coli LPS (1:200, Abcam, ab35654), Mouse Anti-Gram Positive Bacteria antibody (1:100, Abcam, ab20344), Rat Anti-mouse F4/80 Clone BM8 (1:100, Biolegend #123102), Pig Anti-Cytokeratin 8+18 (1:400, Abcam, ab194130), Purified anti-mouse CD31 Antibody (1:100, Biolegend #102501); Anti-Collagen I antibody (1:100; Abcam #254113), Purified anti-mouse CD31 Antibody (1:100, Biolegend #102501) overnight at 4 °C, followed by staining with fluorophore-conjugated secondary antibodies: Donkey Anti-Mouse IgG H&L conjugated Cy5 (715-175-150, Jackson ImmunoResearch); Donkey Anti-Pig IgG H&L 488

(706-545-148, Jackson Immunoresearch) ; Donkey Anti-Rabbit 488 (711-545-152, Jackson Immunoresearch) in blocking buffer for 1h at room temperature. Slides were then stained with DAPI (1 $\mu$ g/mL (1x), sigma) for 5min, and mounted with Antifade Mounting Medium (Beyotime #P0126).

### Immunohistochemistry assays

Immunohistochemistry staining was performed according to standard method, including a deparaffinization and rehydration step. Antigens were retrieved in slight boiling citrate buffer (10mM Sodium Citrate, 0.05% Tween 20, pH6.0) for 15min in microwave at low-to-medium power. Endogenous catalase was quenched by 3% hydrogen peroxide (Propyl alcohol: 30% hydrogen peroxide=9:1) for 10 minutes. After one hour blocking with blocking buffer (NEOBIO SCIENCE, Catalogue ENS004.120), samples were stained with primary Mouse Anti-E. coli LPS antibody (1:100, Abcam #ab20344) overnight at 4 °C in humid chamber followed by half an hour incubation of secondary antibody (HRP conjugated anti-mouse) at RT in humid chamber. Samples were developed using DAB (50mg DAB was dissolved in 0.01m PH7.6 100mL Tris-HCl. BOSTER, Catalogue AR1000) with 1/1000 of 30% hydrogen peroxide. Reactions were terminated by washing with water. Stain with hematoxylin for 10 minutes, wash under running water for 3 minutes, 1% hydrochloric acid (70% ethanol with 1% hydrochloric acid) solution for 3 seconds, wash under running water for 3 minutes then do dehydration, sealed with neutral resin. Blocking buffer and HRP peroxidase conjugated secondary antibody were from Mouse Primary Histostain-Plus IHC Kit (NEOBIO SCIENCE, Catalogue ENS004.120).

### Sample Preparation for Transmission Electron Microscopy

Samples were fixed in 2% paraformaldehyde (16% paraformaldehyde, Ted Pella Co. USA) with 2.5% glutaraldehyde (25% glutaraldehyde ampules, Ted Pella Co. USA) in 0.1M phosphate buffer (pH = 7.3) for 0.5 hour at room temperature, and kept overnight at 4°C. Samples were washed for three times with 0.1M PB (pH 7.3) on ice, and then fixed in 1% osmic tetroxide (4%OsO<sub>4</sub> ampules, Ted Pella Co.USA) in 0.1M PB on ice for 1 hour. Fixed samples were then wash three times with 0.1M PB and three times with ddH<sub>2</sub>O on ice. Tertiary fix with 1% uranyl acetate (UA, Ted Pella Co. USA) in ddH<sub>2</sub>O for 1 hour on ice. After 3X wash by water, samples were dehydrated with 30%, 50%, 70%, 90% ethanol on ice and 100% ethanol at RT. Samples were then infiltrated in acetone (UA, Ted Pella Co. USA): Epon812 (Ted Pella Co.USA)=2:1, acetone: resin=1:2 at RT, each for 30 minutes, followed by 100% Epon812 overnight. Change fresh Epon812 every 3 hours for three times. Samples were then embedded in resin at 60°C for 48 hours. For heavy metal staining, the grids were rinsed briefly in distilled water and post stained for 25 min in 2% (w/v) aqueous uranyl acetate followed by 5 min in 1% (w/v) lead solution. Evaporate enough carbon (5nm) with the dry grids. Images were acquired on 80kV/120kV in a FEI Talos 120kv transmission electron microscopy.

### Flow Cytometric Enumeration of Faecal Bacterial Loads

The experiments were performed as outlined in a published paper (Vandeputte et al., 2017) with minor modifications. Briefly, immediately after opening the cecum, approximately 0.2g of contents were collected in sterile specimen containers and weighed. The contents were resuspended in 1 mL PBS supplemented with 0.05% Tween-20 (to disrupt aggregates of bacteria) (Lehar et al., 2015) and then shaken for 2 min. Centrifuge at 130 g for 5 min to remove fecal particles and then the supernatant was filtered by 40  $\mu$ m cell strainer (BD FALCON) to remove particles. Centrifuge at 1000g for 3min to pellet the cells, 1mL PBS to resuspend the bacteria, and 1mL 70% ethanol as negative control. All the supernatants (1 mL) with 10,000-time dilution were stained with SYTO 9 and propidium iodide (LIVE/DEAD <sup>®</sup>BacLight™ Bacterial Viability and Counting Kit, Invitrogen L34856), and then incubated for 15min at room temperature protected from light. In order to count the absolute number of bacteria in samples, the stained supernatants were added with beads supplied in the kit. The flow cytometry (FCM) analyses were carried out on the CytoFLEX. Where necessary, samples were diluted after staining so that the concentration measured in the flow cytometer was always less than 3 $\times$ 10<sup>5</sup> counts/mL. A threshold value of FSC SSC was applied at 500 and all samples were collected as logarithmic signals.

### Flow Cytometry Profiling of Tumor Immune Cells

Single-cell suspensions from tumors were stained with antibodies with the following antibodies: V450-anti-mouse CD45 (1:200, BD Cat#560501), PerCP/Cy 5.5 antimouse CD3 $\epsilon$ Clone 145-2c11 (1:200, Biolegend Cat#100328), APC anti-mouse CD49b (1:200, Biolegend Cat#103511), PE/Cy7 anti-mouse NK-1.1 Antibody (1:200, Biolegend Cat#108714), APC anti-mouse CD19 Antibody (1:200, Biolegend Cat#115512), APC/Cyanine7 anti-mouse CD11c Antibody (1:200, Biolegend Cat#100722), APC anti-mouse MHCII (1:200, Biolegend Cat#107614), FITC anti-Ly6C (1:200, Biolegend Cat#128006), PE anti-mouse Ly-6G (1:200, Biolegend Cat#127608), PerCP/Cy 5.5 antimouse F4/80 clone BM8 (1:200, Biolegend Cat#123128), APC/Cy7 anti-mouse/human CD11b (1:200, Biolegend Cat#101226). Data were acquired with CytoFLEX and analyzed using FlowJo. The analysis gate was set on the basis of isotype control.

### Circulating Tumor Cell Analysis

Blood was harvested from the right atrium of the mouse using 1-mL EDTA treated syringe (XH0139) and transferred to 2 mL K2EDTA tubes to prevent clotting. Red blood cells were lysed twice by two volumes of ACK lysis buffer (C3702-120mL, Beyotime biotechnology) on ice for 5 min. Cells were then pelleted at 500 x g for 5 min. The cell pellets were resuspended in PBS (C20012500CP, Invitrogen) and 106 cells were spread onto glass slides by cytospin (A78300003, ThermoFisher) at 350 rpm for 5 min. The slides were

then fixed by 4% PFA at RT for 30 minutes, followed by 3 X 3min TBS washes (20X TBS were diluted in DEPC treated water). Samples were then blocked by blocking buffer (TBS in DEPC water+2% BSA+5% Donkey serum+0.1% Triton X-100) at 37 °C for 30 min. Samples were stained with Anti-Cytokeratin 8+18 (1:200, Abcam ab194130) at 37 °C for 1 h. After 3 X 5min TBS washes at RT, samples were incubated with fluorophore-conjugated secondary antibody Donkey Anti-Pig IgG H&L (488) (706-545-148, Jackson ImmunoResearch) at RT for 30 min. To detect the bacteria, samples were then washed by TBS for 3 X 3min, fixed with 4% PFA for 30 minutes at RT, and washed with 2X SSC buffer for 10 min. To completely release bacteria DNA, samples were treated with 1 mg/mL lysozyme (Sigma #62970) and 0.05 mg/mL lysostaphin (Sigma #L7386) for 10 min at RT, respectively. Enzymes were washed off twice by 2X SSC buffer at RT for 5 min and twice by wash buffer for 5 min at RT. After washing, 1 ng/μl probes were hybridized to the cells overnight at 37 °C in hybridization buffer. Unbound probes were wash off by wash buffer at 37 °C for 30min and stained with 1 ng/μl DAPI for 5 min. Slides were then mounted with Antifade Mounting Medium (Beyotime #P0126). Samples were imaged at 400X magnification using an inverted epifluorescence microscope (Nikon Eclipse Ti2). Exposure time was 500ms for the Cy5 fluorescence channel, 50ms for the DAPI fluorescence channel.

### Genetic Labeling of *Staphylococcus xylosus*

To genetically label *Staphylococcus xylosus* and trace the in vivo migration, we inserted an erm (erythromycin resistance)-GFP cassette at the L-lactate dehydrogenase genome locus via recombination-based gene replacement (Brückner, 1997). We cloned the upstream and downstream genomic sequence (around 800 bp) of L-lactate dehydrogenase as the left and right recombination arm. Left arm-erm-GFP-right arm was then constructed into pBT2. Plasmid was max prepped in *E. coli*, and electroporated into *S. xylosus* with up to 5 μg of plasmid DNA pulsed at 2 kV, 25 μf, 100 Ohm. The cells were outgrown and spread onto BHI agar plate containing 10 μg/ml chloramphenicol, and cultured until colonies appeared. Colonies were inoculated into B-medium containing 10 μg/ml erythromycin at 30 °C for recombination. The plasmid was subsequently eliminated by culturing at 40 °C with 2.5 μg/ml erythromycin resistance for several rounds (Brückner, 1997). Bacteria clones devoid of plasmid but with successful genome integration are sensitive to chloramphenicol resistance, but resistant to erythromycin. Therefore, colonies that can grow on erythromycin plate but not chloramphenicol plate were candidates for successful recombination. Positive recombinant clones were genotyped (Forward primer: AGGTATTGTGGTGATTGCA. Reverse Primer: TCACTCGTAAAAAGTTTTGAGA) (Table S6) and sequenced to confirm successful genetic integration.

### Bacteria Tracing Experiment

MMTV-PyMT mTmG mice were used for tracing experiment. When the tumors were palpable, clear all the bacteria with doxycycline in drinking water (0.2 mg/ml) for 5 days. 10<sup>7</sup> CFU of recombinant *S. xylosus* were injected into tumor three days after doxycycline withdrawal, once every 10 days for 3 times. Seven days after the third injection, mice were sacrificed to characterize lung metastasis and lung microbiota by culture. All the homogenate of each sample was spread on CBA plates, split into 10 plates for each mouse. The apparent *S. xylosus* colonies were picked out to perform PCR and genotyped for the recombinant strain. For IV injection assay, 10<sup>7</sup>/mouse recombinant *S. xylosus* were injected via tail vein into mice, once every 10 days for three times, and seven days after the last injection, the lungs were dissected and analyzed for bacteria clones on the culture plate. Meanwhile, 10<sup>6</sup>/mouse PyMT tumor cells, which were cocultured with recombinant *S. xylosus* (invasion rate about 1-5%, the actually bacteria associated tumor cells was estimated to be 10<sup>4</sup>/mouse), were injected through tail vein for one shot, and the lungs were analyzed for recombinant *S. xylosus* clones seven days after the injection.

### Cell viability test and F-actin staining

To test the cell viability after bacteria invasion of tumor cells, we cocultured tumor cells with CFSE-labeled designated bacteria strains and sorted 1.2X10<sup>4</sup>/well bacteria-invaded tumor cells into 24 well-plate with the colony forming medium in the absence of Y27632 (Sigma). Eight hours after attachment, we harvested tumors cells both in the supernatant and adhering on the plate, and stained with Annexin V (Biolegend, #640912) following manufacturer's instructions. Cells were then analyzed by FACS using BC CytoFLEX LX (Beckman) and quantified by FlowJo software (Tree Star, Inc., Ashland, OR).

To visualize actin cytoskeleton, the attached tumor cells were fixed by 4% paraformaldehyde 15 min at room temperature, and stained by Alexa Fluor 488-phalloidin (1:500, Invitrogen #A12379) for 1h at room temperature. Bacteria were stained by vancomycin-Cy5 (4 μM) (Wang et al., 2019) at room temperature for 30 min. Images were acquired on a Nikon Eclipse Ti2 microscope (Nikon). The average actin filament fluorescence intensity was quantified by ImageJ, using the total fluorescence subtracted of the background and divided by the pixel number.

### Microfluidic Circulatory System

Our microfluidic circulatory system with a peristaltic pump (SHENCHEN #LabS3) was assembled according to the published literature (Regmi et al., 2017). Our microfluidic system aims to mimic the fluid shear stress of human bloodstream by controlling the flow rate. The Poiseuille's equation is used to calculate the shear stress that cells experienced.

$$\tau = \frac{4Q\eta}{\pi R^3}$$

where  $Q$  is the flow rate in ml/sec,  $\eta$  is the dynamic viscosity of the fluid, which equals to 0.012 dynes.sec/cm<sup>2</sup>, and  $R$  is the inner radius of the circulatory tube which is 0.15mm. Here we tested different flow rates and calculated corresponding shear stress and chose two condition to do cell viability test. Specifically, when  $Q$  equals to 0.85ml/60s, the shear stress is approximately 14 dynes/cm<sup>2</sup>; when  $Q$  equals to 1.2ml/60s, the shear stress is approximately 20 dynes/cm<sup>2</sup>.

Before experiment, whole circulation system was sterilized by washing with 75% ethanol, followed by ddH<sub>2</sub>O, finally DMEM/F12 (#11320-033, Gibco). We have tested the cell density and circulating time and found  $2.5 \times 10^4$  cells/ml is appropriate for collecting cells after 0.5h circulation in this system. Then the fluid stressed cells were plated onto culture dish for 8h, cell viability test was subsequently performed as mentioned above.

### Rho activity measurement by FRET

The single-chain biosensor with intramolecular fluorescence resonance energy transfer (FRET) that responds to RhoA activation was used to determine RhoA activity according to published protocol (Pertz et al., 2006). The PCDH lentivector backbone was used for the cloning of the biosensor and the complete biosensor can be amplified from pTriEx-RhoA FLARE.sc Biosensor WT (Addgene plasmid # 12150) with forward primer: GGACTAGTcgttacataacttacgtaaat, reverse primer: ACGCGTCGACatgatgcttccgagtgaga (Table S6), and ligated into PCDH-EF1 lentivector (linearized by SpeI and Sall). Then lentivirus was packaged using 293T cells. MMTV-PyMT cells were transduced with lentivirus. 48 hours after transduction, cells were cocultured with bacteria (stained with CellTrace™ Far Red Cell Proliferation staining, #C34564) for 12 hours and bacteria invaded cells were sorted for plating. 8 hours after plating, wash cells once with PBS and fix cells for 30 minutes at room temperature with 4% PFA. Wash cells with PBS for three times. Images were obtained using Nikon A1R HD25. For emission ratio imaging, the following filter sets were used: CFP: 482/35, YFP: 540/30, FRET: 540/30. FRET images were taken before bleach. CFP and YFP images were taken before and after bleach. The YFP intensity of the bleach region was used to confirm the bleach efficiency which should be more than 90%. CFP intensity was used to calculate FRET ratio.  $\text{FRET ratio} = \frac{I_{\text{CorrAfter}} - I_{\text{Before}}}{I_{\text{CorrAfter}}}$ . The background correction was done by subtracting the mean intensity of an ROI in the image where there were no cells. The nonspecific bleaching correction was done by measuring the intensity loss of an ROI in the other part of the cell.  $I_{\text{CorrAfter}} = I_{\text{After}} * \frac{I_{\text{OtherBefore}}}{I_{\text{OtherAfter}}}$ .

### Rho activity measurement by pull down assay

For RhoA-GTP pull down assay, we co cultured PyMT mTmG tumor cells with bacteria, and sorted  $1 \times 10^6$  cells, then performed pull down assay following instructions of Active Rho Detection Kit protocol (CST #8820S). Cells were lysed by ice-cold 1X Lysis/Binding/Wash Buffer plus Aprotinin (Sigma #A1153-25MG), Pepstatin A (Sigma # P8-25MG) and incubated on ice for 5min. The lysates were centrifuged at 16,000g for 15 min at 4°C. 100ul of the supernatants were supplemented with 8ul GST-RBD and incubated for 30min on ice. Then 20ul GSH beads were added and rotated for 1hour at 4°C. Samples were washed three times in 1X Lysis/Binding/Wash Buffer and heated with SDS-PAGE Sample Loading Buffer (Beyotime #P0015L) at 100°C for 10min. Finally, the supernatant was immunoblotted with RhoA antibody. At the same time, whole cell lysates were also immunoblotted for total RhoA as input control.

### Western Blot

Proteins were loading to 4%-20% SurePAGE (GenScript) and immobilized onto PVDF membrane. Western analysis was conducted by blocking the membrane in TBS containing 0.1% Tween-20 and 5% Bovine Serum Albumin (Sigma #V900933-100G) for 1 hour at room temperature. Subsequently, the samples were stained overnight at 4 °C with primary antibodies: Total RhoA (CST #8789, 1:1000), Anti-ROCK2 (phospho S1366) antibody (Abcam # ab228008, 1:1000), beta actin mouse McAb (proteintech # 66009-1-Ig, 1:5000). Then membranes were washed 3X in TBST and followed by addition of secondary antibodies: Anti-Rabbit IgG HRP-Linked antibody (Cell Signaling Technology # 7074P2), Anti-mouse IgG HRP-linked antibody (Cell Signaling Technology # 7076P2). After washing three times, Supersignal Femto Western Blotting substrates (ThermoFisher Scientific) were used and bands were visualized using GEL Imaging System (GE # A1680RGB).

### 3D co-culture system and organoid imaging

Dissociated PyMT tumor cells were then cultured on 2D plate for 2-3 passages until the tumor cells adapted to the in vitro environment. For bacteria and tumor cell coculture, 40uL of growth factor reduced Matrigel (BD Bioscience) was plated on the U-bottom 96 well plate (Costar) and then was solidified at 37 °C for 5 min. 5000 tumor cells were resuspended in 200uL culture media with DMEM/F12 (#11320-033, Gibco) +2%FBS (#SE100-B, VISTECH)+B27 (1X, #12587010, Invitrogen)+PS antibiotics (Gibco) supplemented with EGF (10ng/uL, BD Bioscience), Rspo1(250ng/mL, R&D), ROCK inhibitor Y27632 (10uM, Sigma), and were subsequently overlaid on top of Matrigel. Plate was maintained in 37 °C incubator at 5% CO<sub>2</sub> for 5-6 hours. Bacteria isolated from tumor tissue were cultured in BHI (brain heart infusion) (Solarbio-#LA0360) under in 37°C 200rpm shaker to mid-log phase, then were labelled with CFSE (Invitrogen™ eBioscience™ CFSE, #65-0850-84) following manufacturer's instructions. CFSE labeled bacteria were then added to tumor cells at a ratio of 50:1 in antibiotics free medium. After two days' culture, Matrigel was dissociated with dispase (Sigma) 1mg/mL for one hour to release colonies. The isolated organoids were fixed by formaldehyde and imaged on 3i Marianas Lightsheet Microscope and data were analyzed by Amira (Thermo Fisher).

### Tissue clearing, imaging and analysis

For metastasis colonization assay of clustered tumor cells, FVB female mice at 6–8w were injected with 200 organoids/mouse with and without invaded bacteria through tail vein. The organoid number was counted under the microscope. For metastasis colonization assay of individual cells, dissociated individual tumor cells with/without invaded bacteria were FACS sorted, and  $1 \times 10^4$  cells/mouse were injected into NPSG mice through the tail vein. Lungs were harvested and perfused 2 months after injection and were subsequently fixed in 4% Paraformaldehyde overnight at 4°C on a shaker. After 3–5X PBS washes, samples were cleared in CUBIC1 (a mixture of 25% urea (Sigma), 25% Tetrakis (Sigma), 15% Triton X-100 (Sigma) and dH<sub>2</sub>O) followed by CUBIC2 reagents (a mixture of 25% urea (Sigma), 50% sucrose (Sigma), 10% triethanolamine (Sigma) and dH<sub>2</sub>O) according to published protocol (Susaki et al., 2015). First, samples are treated with 1/2 CUBIC1-H<sub>2</sub>O on a shaker at 37°C overnight. Then PBS wash 3–5 times, each for one hour at RT on a shaker. Subsequently, treated with CUBIC1 3–5 times on a shaker at 37°C per day. Then treated with 1/2 CUBIC2-PBS on a shaker at 37°C overnight. Treated with CUBIC-2 3–5 times on a shaker at 37°C per day. Tissue can be kept in the CUBIC2 until imaging. Cleared samples were imaged on lightsheet Z1 (MC-LM1) (Zeiss). Data were processed and analyzed by Imaris software version 9.6. For analysis of tumor organoids tail vein injection assay, we manually counted colonized tumor foci in each image slices of all the five lobes of mouse lung. For analysis of single cells tail vein injection assay, we quantify the metastatic tumor cells automatically by the “spot” function of Imaris. The screening size threshold is >17 μm and the noise signals outside the lung were excluded.

### Single cell sequencing

For single cell sorting, mTmG PyMT tumor organoids with/without CSFE-labeled bacteria were dissociated into single cells by Dispase (Sigma) (1 mg/ml) for one hour followed by TrypLE (Gibco) for 5 minutes. Single cells with bacteria were sorted by FACS (Moflo, Beckman) into individual well of 96-well PCR plates containing preloaded lysis buffer ERCC spike-in and barcode. Libraries were constructed for 96 control tumor cells with no associated bacteria, 48 tumor cells with associated *E. faecalis*, 96 tumor cells with associated *L. animalis*, 96 tumor cells with associated *S. cuniculi*, 72 tumor cells with associated *S. sanguinis* and 96 tumor cells with associated *S. xylosum*, according to the established protocol (Dong et al., 2018). To lyse the cells, the 96-well plate was first incubated in 72°C for 3 min and then transfer to ice immediately. 2.85 μL of RT mixture containing 40 U SuperScript II reverse transcriptase (Invitrogen, #18,064,071), 5 U RNase Inhibitor (Takara, #2313B), 5× Superscript II first-strand buffer, 25 mM dithiothreitol, 5 M betaine (Sigma-Aldrich, #B0300), 30 mM MgCl<sub>2</sub> (Sigma-Aldrich, #63,020), and 1.75 μM template switch oligo (TSO) primer were added into the lysate. The reverse transcription was conducted at 25°C for 5 min, 42°C for 60 min, 50°C for 30 min, and then 70°C for 10 min. Next, 7.5 μL of PCR mixture containing 6.25 μL 2× KAPA HiFi HotStart ReadyMix (KK2602), 300 nM ISPCR oligo (AAGCAGTGGTATCAACGCAGAGT) (Table S6), and 1 μM 3′ Anchored oligo (GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC) (Table S6) were added to each reaction. The sample was amplified with first initial denaturation at 95°C for 3 min, then 4 cycles of 98°C for 20 s, 65°C for 30 s, and 72°C for 5 min, followed by 10–16 cycles of 98°C for 20 s, 67°C for 15 s, and 72°C for 5 min; and finally 72°C for 5 min. Then the PCR products with different barcodes were pooled together and purified with DNA Clean & Concentrator-5 once (Zymo Research, #D4014), eluted in 50 μL of H<sub>2</sub>O following 0.8× XP beads (Beckman, A63881 AMPure XP) purification twice, finally eluted in 21 μL H<sub>2</sub>O. Next, the cDNAs were amplified with biotinylated index primer (Biotin/CAAGCAGAA GACGGCATAACGAGATindexGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC) (Table S6) and ISPCR oligo for an additional four to five cycles following purification with 0.8× Ampure XP beads again. The biotinylated cDNAs were sonicated (COVARIS #SIA-UH006) into approximately 300~bp fragments. To enrich the amplified products, Dynabeads MyOne Streptavidin C1 Beads (Thermo Fisher Scientific, #5002) were used following manufacturer's instructions. Libraries were prepared using KAPA Hyper Prep Kits (KK8505) and did end repair, A-tailing and adapter ligation by using NEB U-shaped adapter. After post-ligation cleanup, libraries were amplified 7–8 cycles following purified with 0.8× Ampure XP beads twice and eluted in 30 μL of H<sub>2</sub>O. Quality was checked by Fragment Analyzer-12/96 (GENE-QC006). Finally, the libraries were sequenced on a Novaseq platform to generate 150-bp paired-end reads (sequenced by Novogene).

### Single cell RNA-seq data analysis

Raw sequencing reads in FASTQ files were trimmed using TrimGalore (<https://github.com/FelixKrueger/TrimGalore>), where quality trimming was carried out ( $q < 20$ ). Then use UMI-tools to get the count matrix. Briefly, we get the top 100 barcode list from trimmed reads using whitelist tool in UMI-tools (Smith et al., 2017). Then add barcode and UMI information extracted from read2 to read1 using extract tool in UMI-tools. After that, read1 were aligned to the mm10 genome using STAR (Dobin et al., 2013) with default parameter except that the outFilterMultimapNmax was set to 1. Finally, we got the count matrix using featurecounts and count tools in UMI-tools for all downstream analysis. We only kept cells with barcodes that we used when building library. Then filter out cells with ERCC percentage larger than 5%. We only kept cells with genes number larger than 1000 and less than 8000. For gene filtering, we kept genes expressed in more than 10 cells.

The final dataset contained 483 cells with a median of 55232 UMIs and 5023 genes, including 91 control cells, 45 cells with *E. faecalis*, 94 cells with *L. animalis*, 93 cells with *S. cuniculi*, 70 cells with *S. sanguinis* and 90 cells with *S. xylosum*. We used the Seurat (Butler et al., 2018) R package to carry out data normalization and scaling, as well as downstream dimensionality reduction, clustering, tSNE plot overlaying and differential expression. Briefly, we normalized UMI counts by CPM to account for differences in coverage across cells, except that the scaling factor we used was 10000, and then taking the log. For scaling, we regressed out

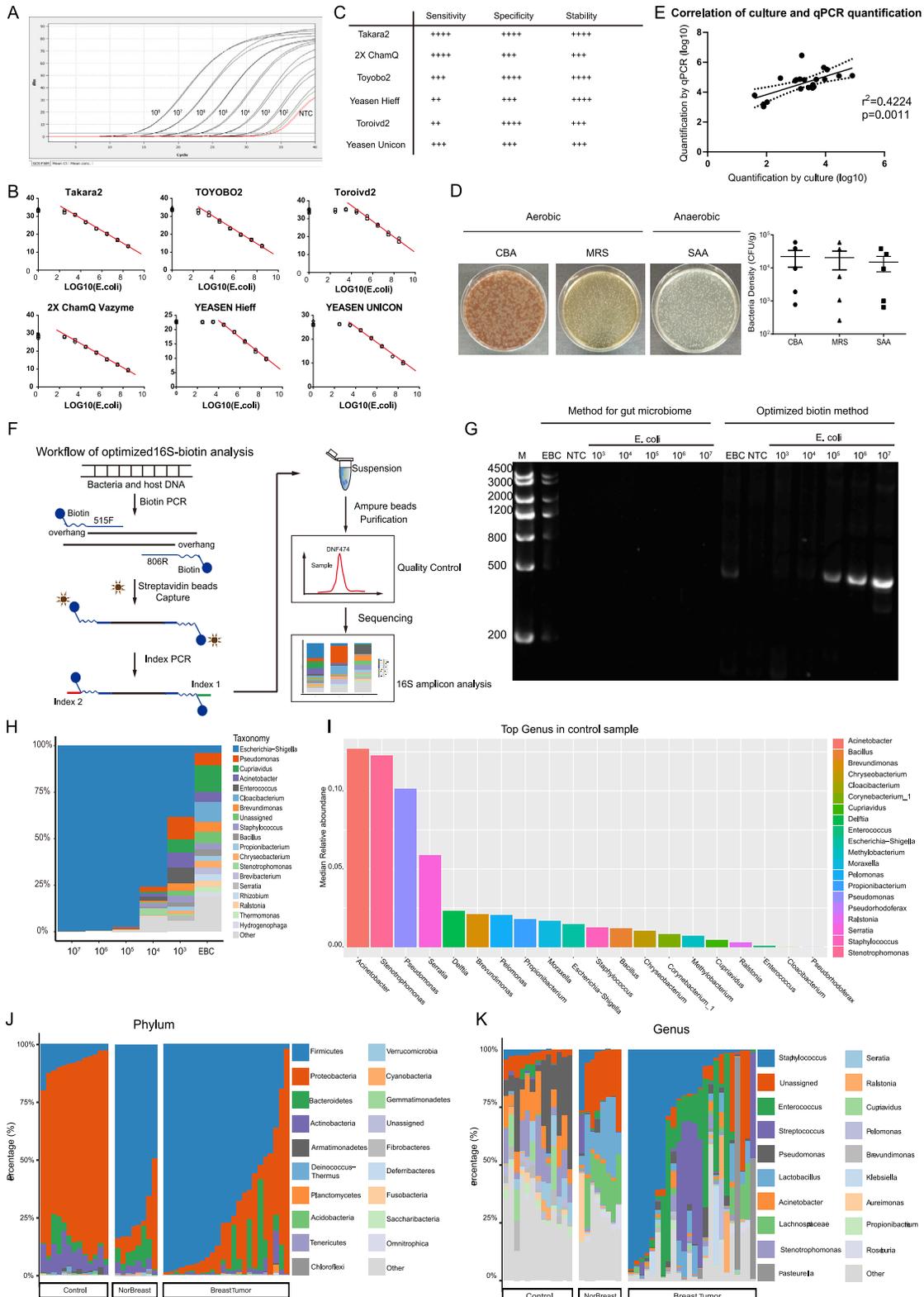
the cell cycle effect, ERCC percentage effect and gene number effect by using `vars.to.regress` parameter. For dimension reduction, we first ran `RunPCA` function then used the first 10 principle component to run `RunTSNE`. Then we got the cluster information using `FindNeighbors` and `FindClusters` functions with original Louvain algorithm and the resolution was set to 1.

For differential gene analysis and pathway enrichment analysis, we compared every treated group with control group using `FindMarkers` function in Seurat (Butler et al., 2018). We identified genes that are expressed in at least 10% cells and the least average expressing  $\log(\text{foldchange})$  is 0.25 and FDR is less than 0.05 as significantly differentially expressed genes. The testing method we used was Wilcoxon rank sum test. Then we separated genes into two parts. One was higher in the group with bacteria, the other was higher in control group. Then we did pathway analysis using `clusterProfiler` R package. We identified pathway with  $\text{FDR} < 0.05$  as significant pathways. The expression of some marker genes we showed was plotted using `FeaturePlot` and `VlnPlot` function in Seurat (Butler et al., 2018) R package. The pathways we showed in dot plot was plotted using `ggplot2` (Wickham, 2016) R package.

### QUANTIFICATION AND STATISTICAL ANALYSIS

Asterisks in the figures indicate the level of statistical significance (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , and \*\*\*\* $p < 0.0001$ ) as determined using either two-tailed unpaired Student t test or Mann-Whitney test as defined in figure captions. Tests were performed using GraphPad Prism software (Version 8, Graphpad Software, La Jolla, CA, United States). Data are expressed as mean  $\pm$  SEM unless otherwise stated.

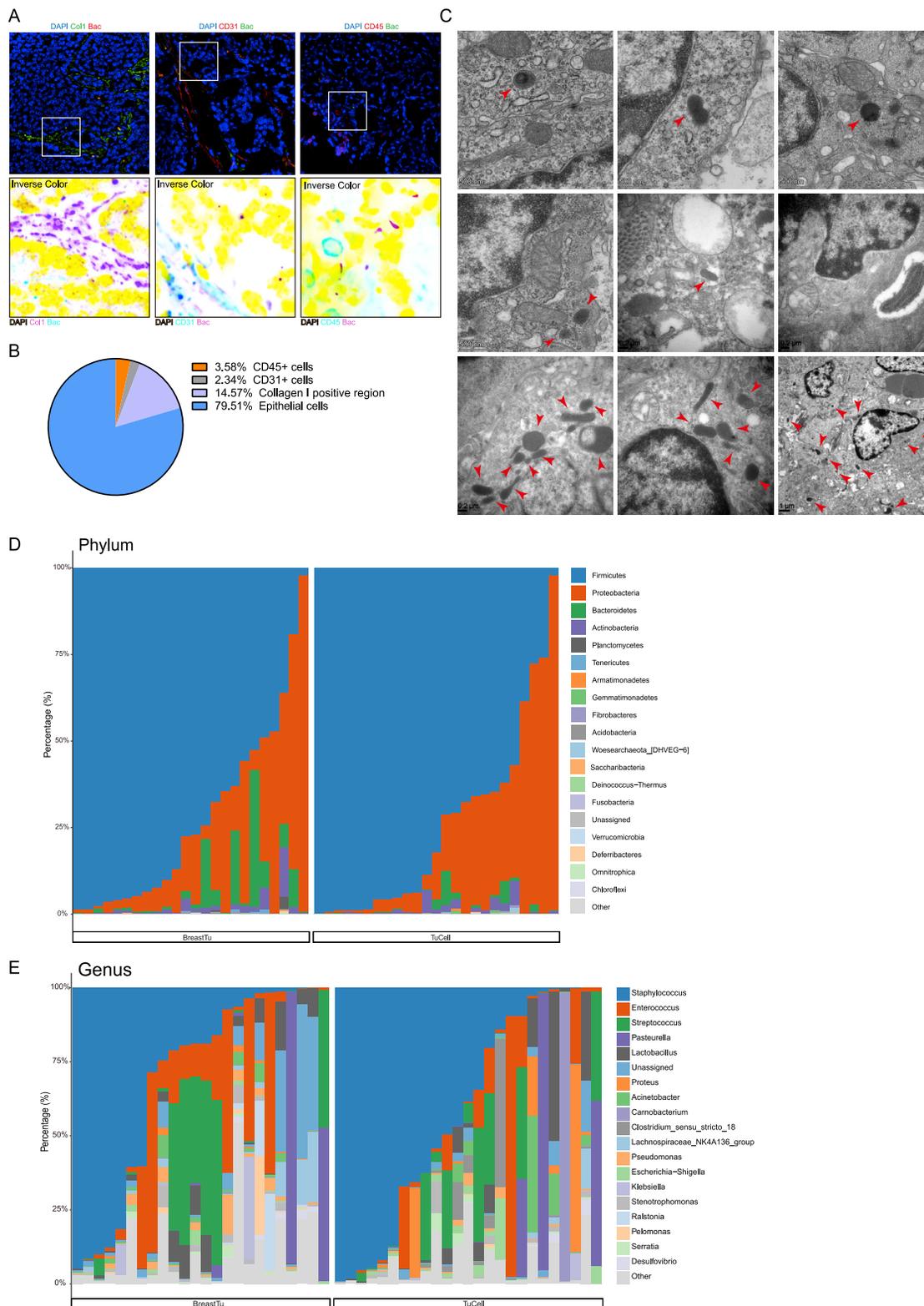
# Supplemental figures



(legend on next page)

**Figure S1. Optimization of analysis of tissue-resident microbiota and relative abundance of microbiota in EBC control, normal breast, and breast tumor at phylum and genus level, related to Figure 1 and STAR Methods**

**A)** Representative real time PCR plot showing the sensitivity of a certain DNA polymerase. **B)** Various commercialized polymerases were tested for their sensitivity, specificity and stability for qPCR quantification of *E. coli*. The sensitivity was determined by the lowest *E. coli* quantity within the lineage range, the specificity was determined by the qPCR method used. Generally, Taqman qPCR is more specific than Sybergreen PCR. The stability was determined by the variation between different experiments. **C)** Polymerase feature table showing the scoring of the sensitivity, specificity and stability of various enzymes. **D)** Homogenized tissue blurry was spread onto CBA plate and MRS plate for aerobic culture, and SAA plate for anaerobic culture. Lower panel, quantification of the bacteria load in the tumor tissue under different culture conditions. Data are presented as mean $\pm$ S.E.M. **E)** Correlation of bacteria load quantification by culture and qPCR. **F)** Schematic diagram showing the workflow of 16S library construction of tissue resident microbiota. To increase the recovery efficiency of PCR products in first step during library preparation, the overhang of amplification primer was modified by 5' biotin. After the first-round biotin PCR, Dynabeads™ MyOne™ Streptavidin C1 Beads (Thermo Fisher Scientific, #5002) were used to pull down the biotin PCR products following manufacturer's instructions. The Streptavidin beads containing first round PCR product was then directly subjected to second round index PCR. Specifically, dual indices from TruePrep® Index Kit V3 for Illumina® (5ul for each index, Vazyme #TD203) were added to target amplicons in a second round PCR using 2 $\times$  Kapa HiFi Hotstart ReadyMix DNA polymerase. Then the Ampure XP beads (Beckman, A63881 AMPure XP) were added to clean up the final libraries and the quality of these purified libraries was checked by Fragment Analyzer-12/96 (GENE-QC006). Finally, samples were sequenced on the Novaseq with 2X250 bp paired-end reads (Novogene) **G)** Agarose gel showing the library preparation from various amount of *E. coli* using two different methods. **H)** 16S rDNA sequencing showing the relative abundance of microbes in various *E. coli* samples. Note: *E. coli* 10<sup>3</sup> sample showed more than 50% of the sample are contaminating microbes, therefore is not reliable. **I)** Bar graph showing the constitution of contaminating microbes in the environment control sample. **J)** Stacked plot of relative abundance of bacteria at the phylum level in environment control, normal breast and PyMT breast tumor. **K)** Stacked plot of relative abundance of bacteria at the genus level in environment control, normal breast and PyMT breast tumor.



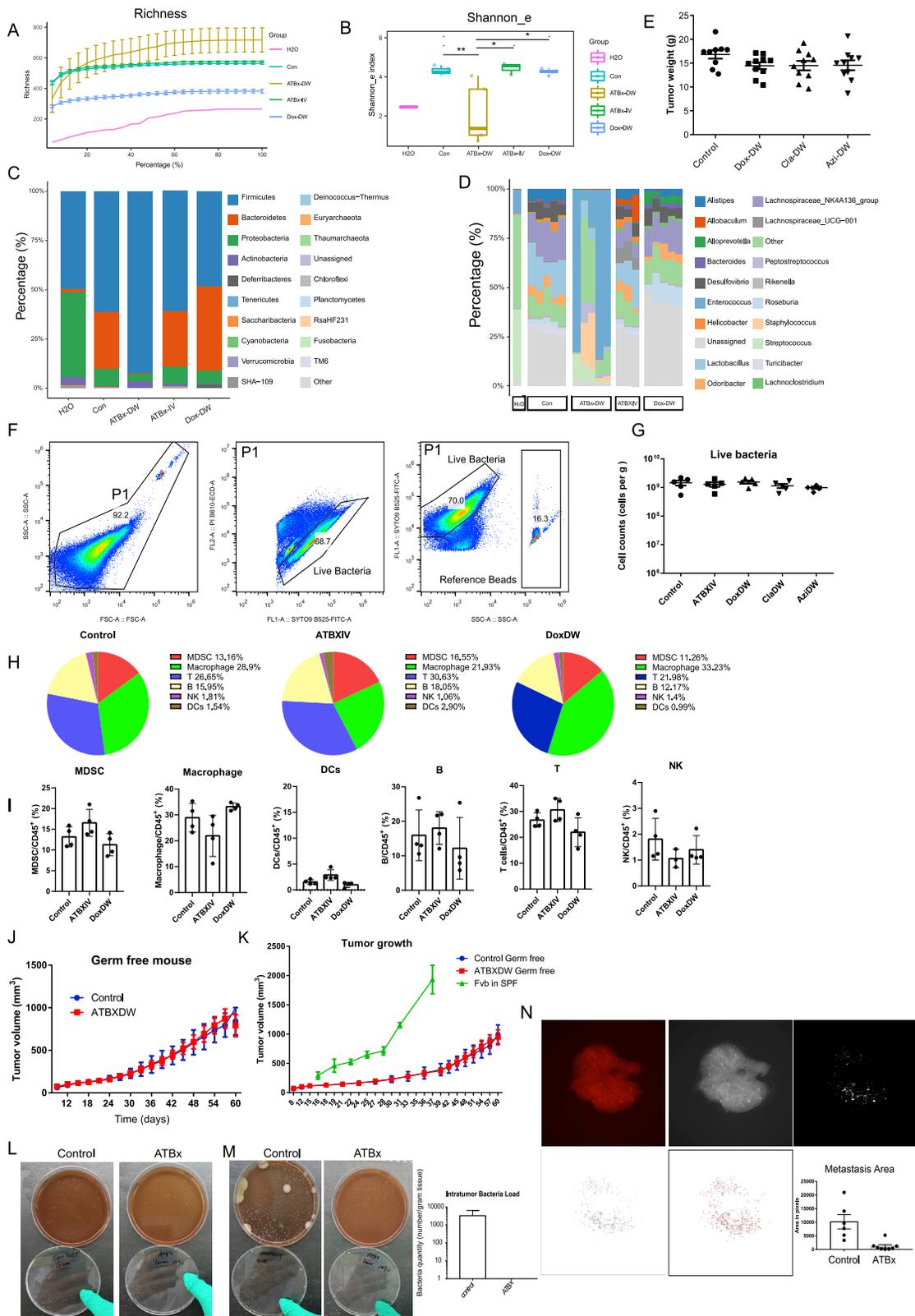
**Figure S2. Bacteria distribution analysis and relative abundance of microbiota in breast tumor tissue, related to Figure 2**

**A)** MMTV-PyMT tumors were sectioned and stained for Collagen I for stromal fibroblasts, CD31 for endothelial cells and CD45 for immune cells. An inverse color images were shown to visualize bacteria vividly. **B)** Quantification of the distribution of bacteria within the PyMT tumor. Note: the majority of bacteria cohabited

(legend continued on next page)

---

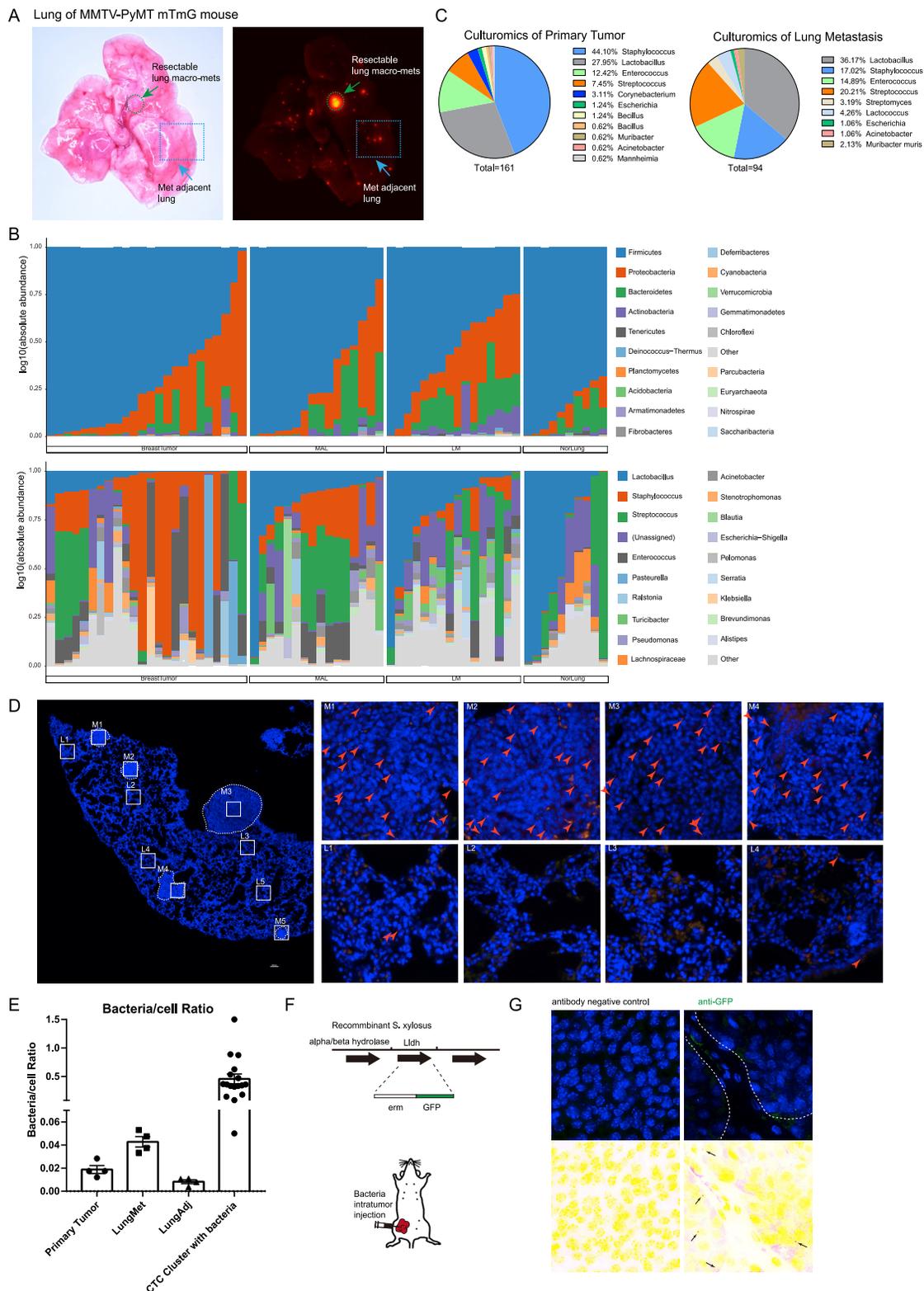
with PyMT tumor cells rather than immune cells, endothelial cells or stromal fibroblasts. **C)** Representative EM images showing bacteria structures within the cytosol of the tumor cells. **D)** Stacked plot of relative abundance of bacteria at the phylum level in PyMT breast tumor tissue and breast tumor cell. **E)** Stacked plot of relative abundance of bacteria at the genus level in PyMT breast tumor tissue and breast tumor cell.



---

**Figure S3. Antibiotics treatment effects on PyMT tumors and gut microbiome, related to Figure 3**

**A)** Richness analysis of gut microbiome taxonomy over the percentage of sampling. **B)** Shannon index analysis showing the alpha diversity of gut microbiome after antibiotics treatment. Wilcoxon test was applied to evaluate the p value. \*  $p < 0.5$ , \*\*  $p < 0.01$ . **C)** Stacked plot of relative abundance of bacteria at the phylum level for gut microbiome treated by different antibiotics **D)** Stacked plot of relative abundance of bacteria at the genus level for gut microbiome treated by different antibiotics. **E)** PyMT spontaneous tumor weight quantification after Doxycycline (0.2 mg/ml), Clarithromycin (0.2 mg/ml), Azithromycin (0.2 mg/ml) treatment for 40 days. Data are presented as mean  $\pm$  S.E.M. **F)** FACS plot showing the quantification of live bacteria in the gut under control and antibiotic treatment condition with SYTO9 and PI staining. **G)** Quantification of the live bacteria in the gut under control and antibiotic treatment condition. Data are presented as mean  $\pm$  S.E.M. **H)** Pie charts showing the composition of various immune cells after control, ATBx-IV treatment and Dox-DW treatment. **I)** Bar graphs showing the percentage of individual cell types in CD45+ immune population. There were no significant changes observed after antibiotic treatments ( $p > 0.05$ ). Data are presented as mean  $\pm$  S.E.M. **J)** Tumor volume of bacteria-containing PyMT-mTmG on germ-free mice in the presence or absence of ATBx drinking water treatment. Data are presented as mean  $\pm$  S.E.M. **K)** Tumor volume of bacteria-containing PyMT-mTmG on germ-free mice in the presence or absence of ATBx drinking water treatment, combined with tumor volume of fvb mice injected with PyMT tumor cells under SPF condition. Data are presented as mean  $\pm$  S.E.M. **L)** Bacteria culture of cecum content from germ free mice under control or ATBx treatment. **M)** Bacteria culture and quantification of tumor tissue from germ free mice under control or ATBx treatment. Note: control tumors contained significant amount of culturable bacteria, while ATBx treated tumors were devoid of bacteria. Data are presented as mean  $\pm$  S.E.M. **N)** metastasis burden analysis by quantification of metastatic area using Image J. Images were converted to grey scale, then segmented the mets by setting threshold, followed with 'particle analysis' function to calculate the total number and area. Data are presented as mean  $\pm$  S.E.M.

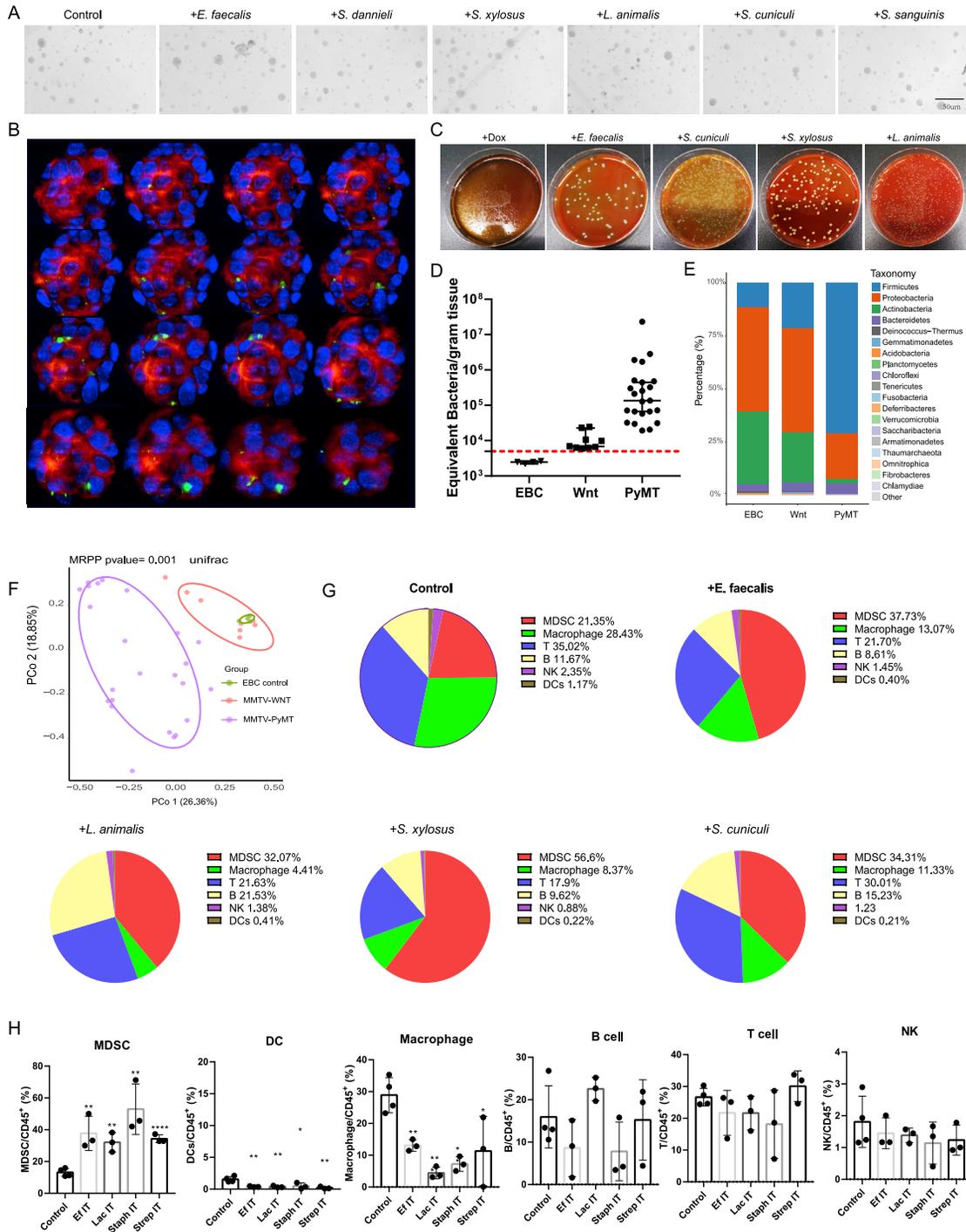


**Figure S4. Analysis of microbiota in breast tumor and lung with metastasis, related to Figure 4**

**A)** bright field image and fluorescent image showing a lung with metastasis in MMTV-PyMT mTmG mouse. Dashed circle showing a resectable metastasis, dashed rectangle showing a met adjacent lung tissue containing many micro-metastases. **B)** Stacked plot of relative abundance of bacteria at the phylum level and genus level in PyMT breast tumor, met adjacent lung tissue, macro-metastasis, and normal lung. **C)** pie chart showing the culture isolated bacteria in primary (legend continued on next page)

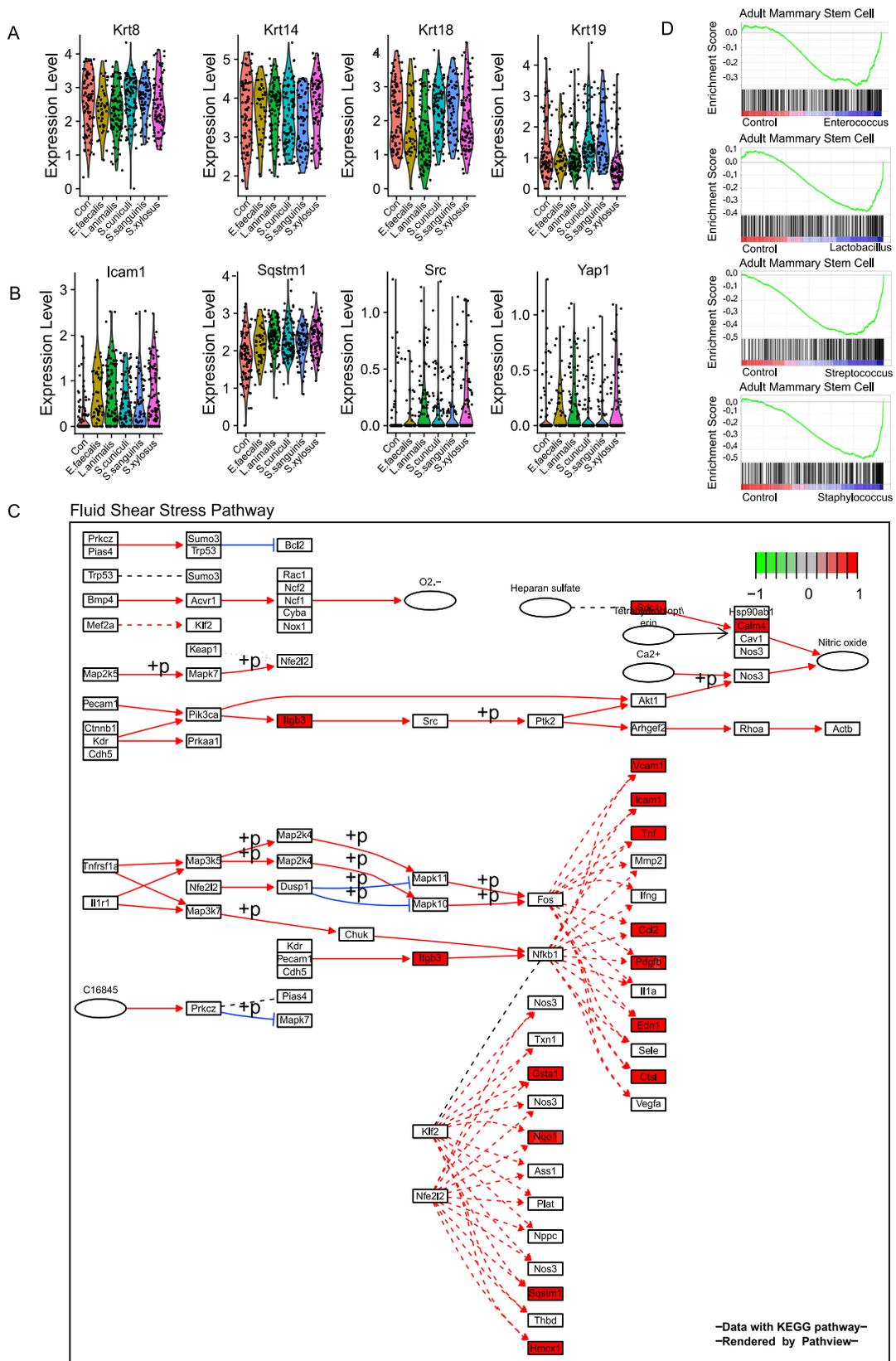
---

tumor and lung metastasis. **D)** 16S FISH image showing the microbiome in a lung tissue with PyMT tumor metastasis. L stands for lung tissue and M stands for metastasis. Red arrowhead pointing to the detected bacteria signal in the designated region. **E)** Quantification of bacteria/cell ratio at primary tumor site, lung metastases, metastasis adjacent lung, and on circulating tumor clusters. Data are presented as mean $\pm$ S.E.M. **F)** Schematic diagram showing the bacteria tracing strategy. **G)** Staining of GFP showing the location of recombinant *S. Xylosus* in the tumor after intratumor administration.



**Figure S5. Bacteria invasion of the PyMT tumor cells, related to Figure 5**

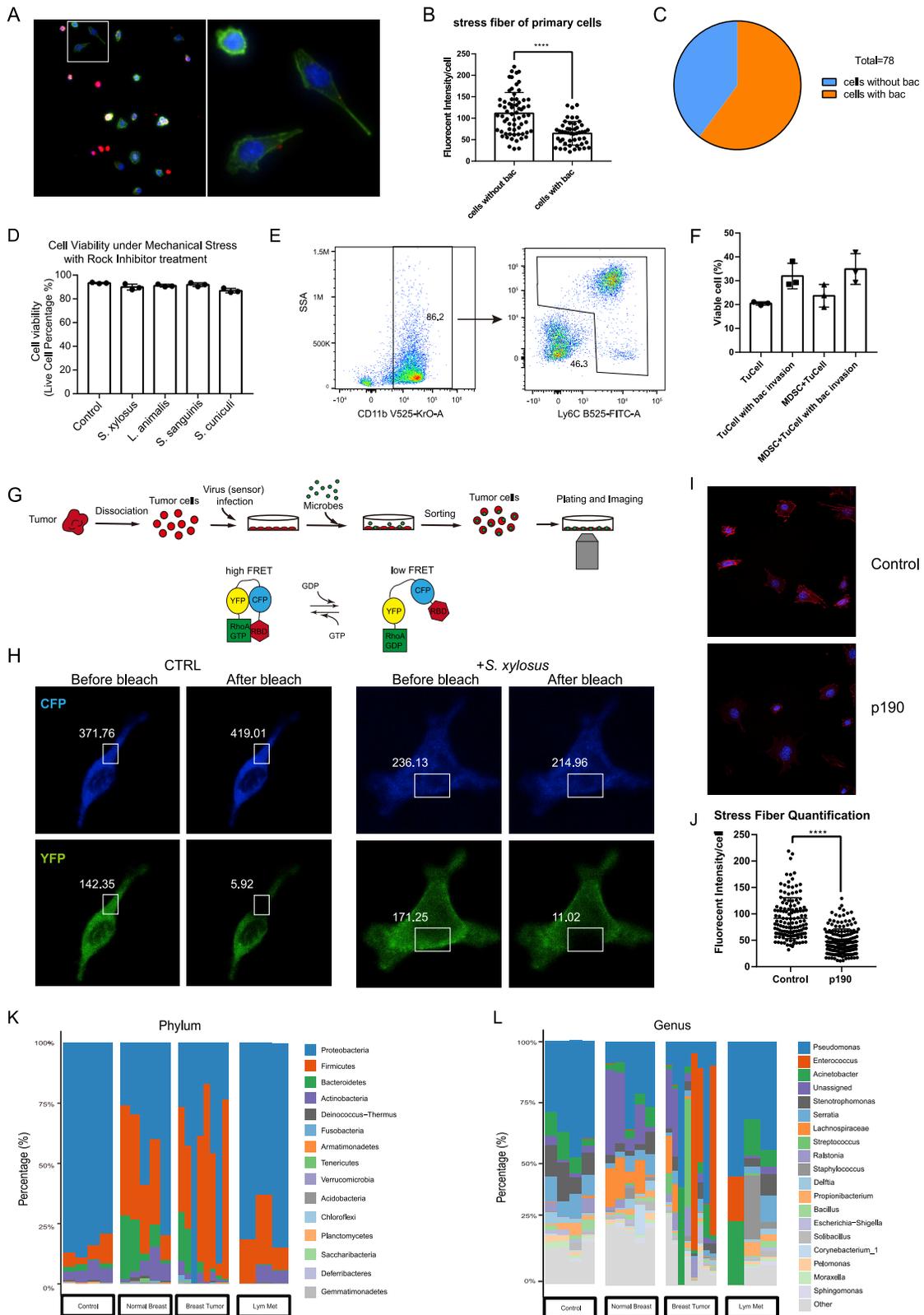
**A)** Coculture experiment showing the effect of various bacteria strains on PyMT tumor cell growth. **B)** optical sections of a representative PyMT organoid invaded by *Staphylococcus xyloso* showing the localization of bacteria at different focal planes. **C)** Culture of intratumor bacteria from PyMT tumors injected with *E. faecalis*, *S. cuniculi*, *S. xyloso*, *L. animalis* Colonies were sequenced and confirmed for the dominance of the injected bacteria strains. **D)** Bacteria quantification by qPCR in Wnt tumors and PyMT tumors. Note the abundance of bacteria in Wnt tumors are markedly lower than that in PyMT tumors. Data are presented as mean±.S.E.M. **E)** Bacteria abundance at Phylum level. **F)** Beta diversity of PyMT tumor and Wnt tumor  $n_{\text{PyMT}}=24$ ,  $n_{\text{Wnt}}=9$ . **G)** Pie charts showing the composition of various immune cells 7 days after intratumor injection of PBS, *E. faecalis*, *S. xyloso*, *L. animalis* and *S. cuniculi*. **H)** Bar graphs showing the percentage of individual cell types in CD45+ immune population 7 days after intratumor injection of PBS, *E. faecalis*, *S. xyloso*, *L. animalis* and *S. cuniculi*. Data are presented as mean±.S.E.M.



---

**Figure S6. Target genes and signaling pathway regulated by bacteria invasion, related to [Figure 6](#)**

**A)** Keratin genes expression after bacteria invasion. **B)** Representative genes involved in fluid shear stress after bacteria invasion. **C)** KEGG pathway plot showing genes regulated by bacteria invasion in the fluid shear stress pathway. **D)** GSEA enrichment analysis showing mammary stem cell program is enriched in bacteria-containing tumor cells.



(legend on next page)

---

**Figure S7. Bacteria's role in primary tumor cells and microbiome in human tissues, related to Figures 6 and 7**

**A)** Tumor cells dissociated from PyMT breast tumor were sorted by Lin-CD49f+EpCAM+ and plated onto 96 well plate. After adhesion, cells were fixed and stained with phalloidin (for stress fiber, green) and cy5-vancomycin probe (for bacteria, red). **B)** Quantification of stress fiber intensity for tumor cells with/without bacteria. **C)** Pie chart showing the quantification of the percentage of cells with/without bacteria. **D)** Quantification of the viability of PyMT cells with/without bacteria invasion after fluid stress in the presence of ROCK inhibitor Y27632 (10 $\mu$ M) treatment. **E)** FACS plot showing the sorting of MDSC cells from bacteria injected tumors. **F)** The survival assay of tumor cells with invaded *S. Xylosus* and MDSC addition. **G)** Schematic diagram showing the FRET analysis of bacteria invaded PyMT cells. PyMT cells were transduced by RhoA FRET sensor and cocultured with designated bacteria species. Bacteria invaded cells were sorted and plated in the imaging plate, subjected for FRET analysis. **H)** Representative demonstration of FRET analysis by photobleaching. Note: Control CFP signal significantly elevated after photobleaching of YFP, indicating a block of energy transfer; in *S. Xylosus* invaded PyMT cells, the elevation of CFP signaling was not observed. **I)** Stress fiber staining by Phalloidin (Cy5) with/without p190/Arhgap35 expression. **J)** Stress fiber quantification by Phalloidin staining with/without p190/Arhgap35 expression. Data were presented as mean $\pm$ S.D. **K)** Relative abundance of microbiota in EBC control, human normal breast, human breast tumor and lymph node metastasis at phylum level. **L)** Relative abundance of microbiota in EBC control, human normal breast, human breast tumor and lymph node metastasis at genus level.