RESEARCH ARTICLE SUMMARY

CANCER MICROBIOME

Commensal microbiota from patients with inflammatory bowel disease produce genotoxic metabolites

Yiyun Cao, Joonseok Oh, Mengzhao Xue, Won Jae Huh, Jiawei Wang, Jaime A. Gonzalez-Hernandez, Tyler A. Rice, Anjelica L. Martin, Deguang Song, Jason M. Crawford, Seth B. Herzon, Noah W. Palm*

INTRODUCTION: Gut microbiota can potentially contribute to the development and progression of colorectal cancer (CRC) by producing small-molecule genotoxins. For example, select commensal *Escherichia coli* strains produce the canonical microbiota-derived genotoxin colibactin, which engenders the formation of DNA double-strand breaks (DSBs) in intestinal epithelial cells and exacerbates CRC in mouse models. Furthermore, human CRCs harbor colibactin-associated mutational signatures, implying a direct role for microbiota-induced DNA damage in CRC. However, the impacts of microbiota-derived genotoxins beyond colibactin remain largely unexplored.

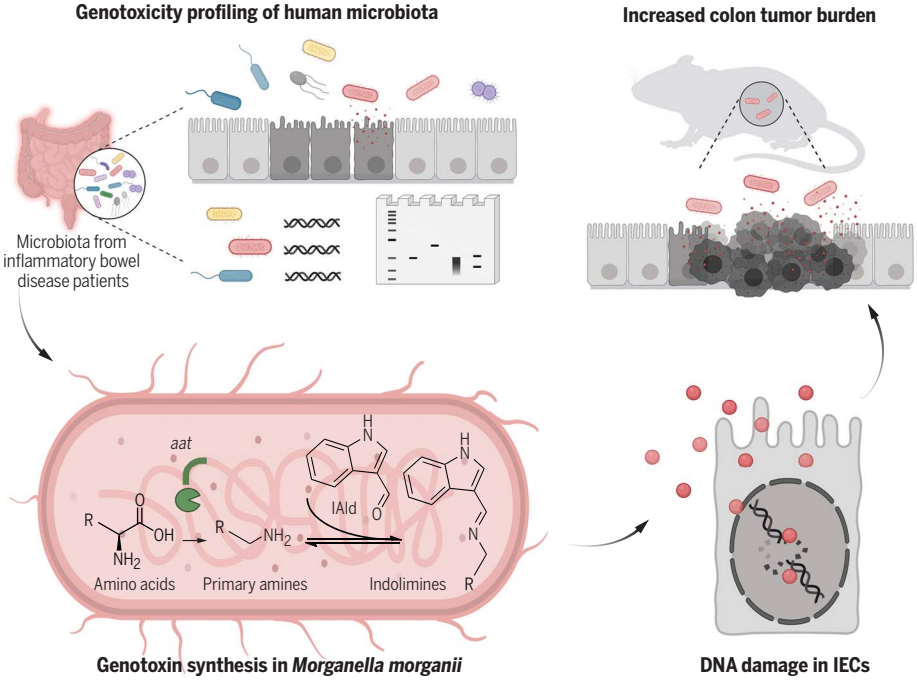
RATIONALE: Given the extensive diversity of small-molecule metabolites produced by bacteria, we hypothesized that additional taxa from the human gut microbiome might produce previously undiscovered small molecules that cause DNA damage in host cells. Identifying and characterizing such genotoxins and their respective biosynthetic pathways may reveal causal roles of gut microbes in shaping host biology and disease susceptibility. Thus, we designed a large-scale electrophoresis-based DNA damage screening pipeline to evaluate the genotoxicity of a collection of more than 100 gut commensals isolated from patients with inflammatory bowel disease (IBD).

RESULTS: We identified diverse bacteria from the human microbiota whose small-molecule metabolites caused genotoxicity in both cell-free and cell-based DNA damage assays. For example, small-molecule metabolites from gram-positive bacteria (including *Clostridium perfringens* and *Clostridium ramosum* strains) and gram-negative bacteria (including multiple *Morganella morganii* strains) directly damaged DNA in cell-free assays and induced the expression of the DSB marker γ-H2AX and cell-cycle arrest in epithelial cells. However, the DNA damage patterns caused by these metabolites were distinct from colibactin-induced cross-links, and these isolates lacked the biosynthetic machinery to produce colibactin or other known genotoxins. These data thus implied the existence of previously unrecognized microbiota-derived genotoxins.

M. morganii is enriched in the gut microbiota of both IBD and CRC patients. Using a combination of comparative metabolomics and bioactivity-guided natural product-discovery techniques, we discovered a family of *M. morganii*-derived small-molecule genotoxins—termed the indolimines—that elicited DNA damage in cell-based and cell-free assays. Furthermore, we identified a previously uncharacterized bacterial decarboxylase (annotated as an aspartate aminotransferase encoded by the *aat* gene) that was essential for indolimine synthesis and constructed an isogenic *aat* mutant *M. morganii* that lacked genotoxicity in both cell-free and cell-based DNA damage assays.

Compared with the non–indolimine-producing mutant, wild-type *M. morganii* caused increased intestinal permeability and induced transcriptional signatures associated with abnormal DNA replication and intestinal epithelial cell proliferation in gnotobiotic mice. Furthermore, indolimine-producing *M. morganii* induced increased colonic tumor burden in the context of a mock microbial community in a mouse model of CRC.

CONCLUSION: By leveraging function-based assessments of the microbiome, we uncovered the existence of a broader universe of microbiota-derived small-molecule genotoxins. We found that diverse bacterial strains isolated from IBD patients exhibited DNA-damaging activities and discovered a previously undescribed family of genotoxins, termed the indolimines, produced by the IBD- and CRC-associated species *M. morganii*. Indolimine-producing *M. morganii* caused increased intestinal permeability and exacerbated colon tumorigenesis in gnotobiotic mice. Overall, these studies imply an expanded role for microbiota-derived genotoxins in shaping host biology and disease susceptibility. ■



Human gut microbes isolated from IBD patients produce small-molecule genotoxins. Diverse gut microbes isolated from patients with IBD exhibit direct genotoxicity. *M. morganii* produces a family of genotoxic small molecule metabolites, termed the indolimines. Indolimine-producing *M. morganii* induces DNA damage in intestinal epithelial cells (IECs) and increased colon tumor burdens in gnotobiotic mouse models.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: noah.palm@yale.edu
Cite this article as Y. Cao *et al.*, *Science* **378**, eabm3233 (2022). DOI: 10.1126/science.abm3233

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abm3233

RESEARCH ARTICLE

CANCER MICROBIOME

Commensal microbiota from patients with inflammatory bowel disease produce genotoxic metabolites

Yiyun Cao[1], Joonseok Oh[2,3], Mengzhao Xue[2]†, Won Jae Huh[4], Jiawei Wang[5], Jaime A. Gonzalez-Hernandez[1], Tyler A. Rice[1], Anjelica L. Martin[1], Deguang Song[1], Jason M. Crawford[2,3,6], Seth B. Herzon[2,7], Noah W. Palm[1]*

Microbiota-derived metabolites that elicit DNA damage can contribute to colorectal cancer (CRC). However, the full spectrum of genotoxic chemicals produced by indigenous gut microbes remains to be defined. We established a pipeline to systematically evaluate the genotoxicity of an extensive collection of gut commensals from inflammatory bowel disease patients. We identified isolates from divergent phylogenies whose metabolites caused DNA damage and discovered a distinctive family of genotoxins—termed the indolimines—produced by the CRC-associated species *Morganella morganii*. A non–indolimine-producing *M. morganii* mutant lacked genotoxicity and failed to exacerbate colon tumorigenesis in mice. These studies reveal the existence of a previously unexplored universe of genotoxic small molecules from the microbiome that may affect host biology in homeostasis and disease.

Colorectal cancer (CRC) is the third most common malignancy and the second leading cause of cancer deaths worldwide (*1*). Two thirds of all CRC cases occur in individuals without a family history of CRC or inherited genetic mutations predisposing an individual to CRC (*2*). Thus, environmental risk factors that promote the acquisition and accumulation of somatic-genetic and epigenetic aberrations are chief contributors to CRC development. The gut microbiome has been reported to modulate intestinal carcinogenesis through diverse mechanisms (*3*–*5*). Examples include short-chain fatty acid–producing *Clostridia* species [which induce regulatory T cells and temper inflammation–induced carcinogenesis (*6*)] and *Fusobacterium nucleatum* strains that enhance tumor growth by inducing epithelial proliferation through FadA-mediated engagement of E-cadherin and activation of Wnt/β-catenin signaling (*7*). Microbial products may also trigger DNA modifications in intestinal epithelial cells (*8*). For example, the 20-kDa *Bacteroides fragilis* toxin induces DNA damage through induction of reactive oxygen species (*9*) whereas cytolethal distending toxin from pathogenic gram-negative bacteria exhibits direct DNase activity (*10*).

Small-molecule metabolites from the microbiome may influence CRC risk by directly damaging DNA. Select *Escherichia coli* strains produce the reactive small-molecule genotoxin colibactin, which alkylates and crosslinks DNA, triggering double-strand DNA breaks (DSBs) that may facilitate intestinal carcinogenesis in mouse models (*11*–*14*). The colibactin biosynthetic machinery is encoded by a 54-kb hybrid polyketide synthase-nonribosomal peptide synthetase (PKS-NRPS) gene cluster referred to as the *pks* or *clb* locus (*11*), and the mature chemical structure of colibactin responsible for the pathway's DNA interstrand crosslinking activity was recently determined (*15*, *16*). Human CRCs also contain mutational signatures consistent with colibactin-induced DNA damage, implicating colibactin in human CRC (*17*, *18*).

The colibactin paradigm illustrates the importance of microbiota metabolite–induced DNA damage in human CRC. However, aside from colibactin the potential role of microbiota-derived small-molecule genotoxins in CRC initiation or progression remains mostly unexplored. Given the substantial complexity and diversity of metabolites produced by bacteria (*19*), we hypothesized that diverse taxa from the human gut microbiome may produce previously undiscovered small molecules that cause DNA damage in intestinal epithelial cells and contribute to the development of CRC. We established a pipeline to evaluate the genotoxicity of small-molecule metabolites derived from more than 100 phylogenetically diverse human gut microbes. We identified a diverse set of microbes that produced genotoxic small-molecule metabolites, including the gram-positive bacteria *Clostridium perfringens* and *Clostridium ramosum* and the gram-negative bacteria *Morganella morganii*. However, none of these isolates produced known genotoxins such as colibactin or encoded known genotoxin-producing biosynthetic gene clusters. We combined untargeted metabolomics and bioactivity-guided natural product discovery techniques to isolate and characterize a family of previously undescribed genotoxic metabolites—termed the indolimines—produced by CRC-associated *M. morganii*. Finally, we decoded the pathway for indolimine synthesis and constructed an isogenic non–indolimine-producing mutant of *M. morganii* that lacked genotoxicity in vitro and in vivo and failed to exacerbate colon tumorigenesis in a mouse model of CRC.

Results

Establishing a pipeline to identify genotoxic gut microbes from patients with inflammatory bowel disease

We established a pipeline to screen diverse human gut microbes based on their ability to directly damage DNA. We then applied this pipeline to a gut microbiota culture collection assembled by anaerobic culturomics of stool samples from 11 inflammatory bowel disease (IBD) patients (*20*), as IBD patients are at a considerably increased risk of developing CRC (*21*, *22*). This collection consists of 122 bacterial isolates that span 5 phyla, 9 classes, 10 orders, and 17 families, as well as multiple strains that were assigned to the same species (Fig. 1A).

To probe for genotoxicity, we evaluated the activity of each isolate in a plasmid DNA damage assay. As genotoxic metabolites such as colibactin can be recalcitrant to isolation (*15*), we focused our primary studies on co-incubation of individual bacterial isolates with linearized pUC19 plasmid DNA (Fig. 1A). This assay is based on the principle that the extent and modes of DNA damage can be assessed by electrophoresis under native and denaturing conditions (*23*): Double-stranded linearized pUC19 DNA migrates as a slow-moving band under native electrophoresis, whereas denaturing treatment separates double-stranded DNA into single-stranded DNA, leading to a band with higher mobility (Fig. 1A, L2). The formation of DNA interstrand cross-links—e.g., by colibactin—prevents unwinding under denaturing conditions, thereby resulting in a band with the same mobility as duplexed DNA (Fig. 1A, L3). Alkylation at many of the sites in DNA is known to decrease the stability of the glycoside bonds, resulting in deglycosylation and fragmentation. These damaged DNA products are detected as smaller fragments of higher mobility

[1]Department of Immunobiology, Yale University School of Medicine, New Haven, CT 06519, USA. [2]Department of Chemistry, Yale University, New Haven, CT 06520, USA. [3]Institute of Biomolecular Design and Discovery, Yale University, West Haven, CT 06516, USA. [4]Department of Pathology, Yale University School of Medicine, New Haven, CT 06510, USA. [5]Program of Computational Biology and Bioinformatics, Yale University, New Haven, CT 06510, USA. [6]Department of Microbial Pathogenesis, Yale University School of Medicine, New Haven, CT 06536, USA. [7]Department of Pharmacology, Yale University School of Medicine, New Haven, CT 06520, USA.
*Corresponding author. Email: noah.palm@yale.edu
†Present address: Laboratory of Genetically Encoded Small Molecules, The Rockefeller University, New York, NY 10065, USA.

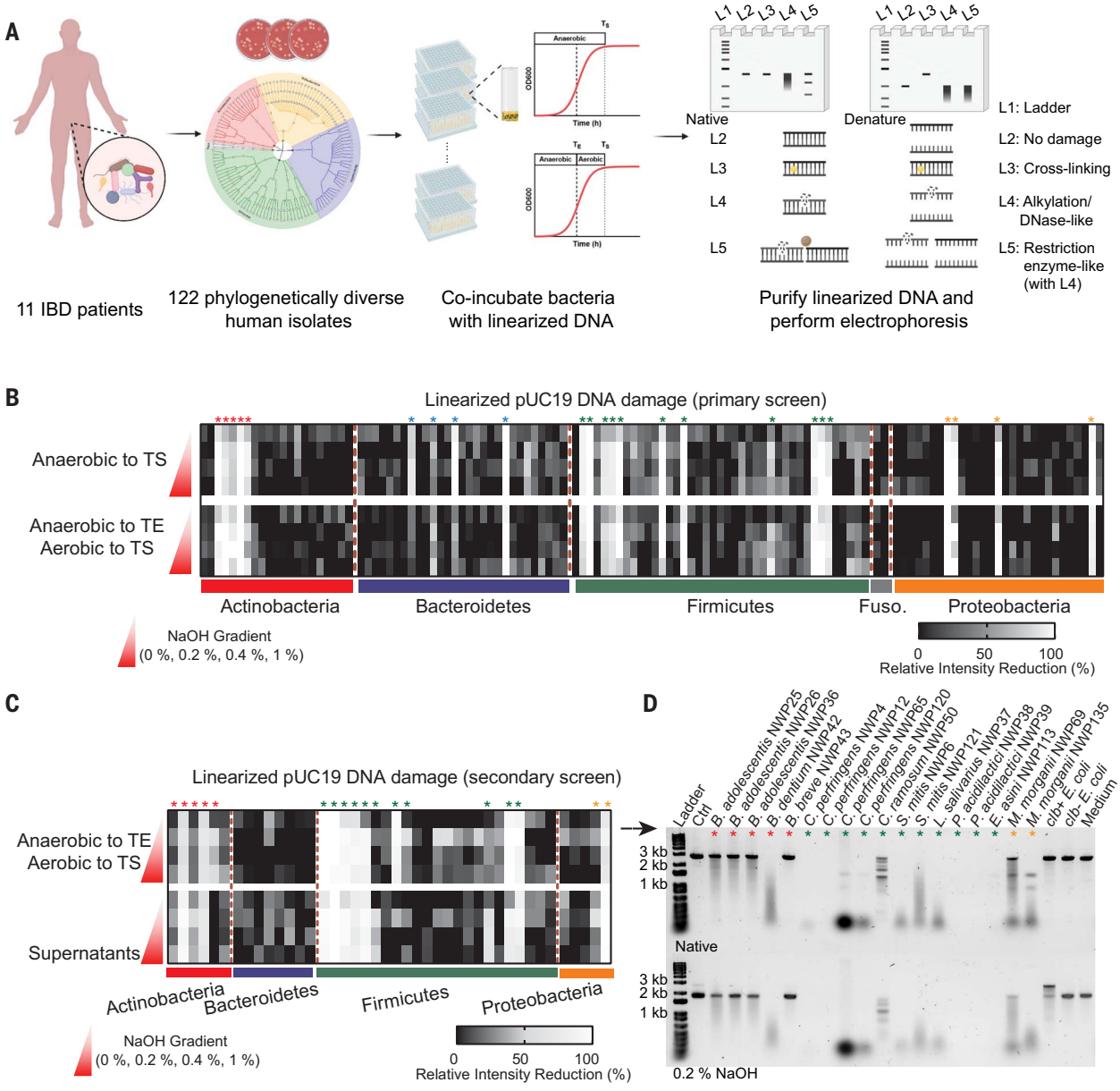


Fig. 1. Establishing a pipeline to identify genotoxic gut microbes from patients with IBD. (**A**) Overview of functional screening of gut microbes for direct genotoxicity; 122 phylogenetically diverse bacterial isolates from 11 IBD patients (shaded based on phylum: red, actinobacteria; blue, bacteroidetes; green, firmicutes; gray, fusobacteria; orange, proteobacteria) were evaluated for genotoxicity through co-incubation with plasmid DNA followed by gel electrophoresis. Bacterial growth curves for all isolates were determined through OD_{600} and individual isolates were cocultured with linearized pUC19 DNA under indicated conditions (T_E, time point of exponential phase; T_S time point of stationary phase). DNA damage was assessed with gel electrophoresis after native or denaturing treatment of purified plasmid DNA. (**B**) Diverse human gut microbes exhibited direct DNA-damaging activities (N = 1 independent experiments). Bacterial genotoxicity was determined by calculating the relative intensity reduction [(RIR) percent] of linearized pUC19 DNA bands after co-incubation with 122 diverse human gut bacteria [as outlined in (A)] under indicated conditions as compared with medium-only controls. Linearized pUC19 DNA was then purified through column purification and treated with or without gradient NaOH (0%, 0.2%, 0.4%, and 1%) before evaluating DNA integrity through gel electrophoresis. The 24 selected putative genotoxic isolates were labeled with asterisks and colors were assigned based on phylum: red, actinobacteria; blue, bacteroidetes; green, firmicutes; orange, proteobacteria. (**C**) Selected isolates consistently exhibited direct DNA-damaging activities under other culture conditions (N = 1). RIR percent of linearized pUC19 DNA bands after co-incubation with live bacteria or supernatants (SUP) of 42 isolates [24 genotoxic isolates selected from (B) and 18 phylogenetically related nongenotoxic isolates] under indicated conditions as compared with medium-only controls. Linearized pUC19 DNA was then purified through column purification and treated with or without gradient NaOH (0%, 0.2%, 0.4%, and 1%) before evaluating DNA integrity through gel electrophoresis. The 18 selected putative genotoxic isolates were labeled with asterisks and colors were assigned based on phylum: red, actinobacteria; green, firmicutes; orange, proteobacteria. (**D**) Representative image of gel electrophoresis for 18 selected genotoxic isolates (N = 2). Linearized pUC19 DNA damage was evaluated after co-incubation with 18 selected genotoxic isolates from (C). Column-purified DNA was treated with or without 0.2% NaOH before gel electrophoresis. Ctrl, nontreated linearized pUC19 DNA; *clb+ E. coli*, colibactin-producing *E. coli*; *clb− E. coli*, non–colibactin-producing *E. coli*; Medium, medium-alone treated linearized pUC19 DNA.

following electrophoresis (*24*) (Fig. 1A, L4). Extensive DNA damage, for example, by DNase-mediated degradation, results in a loss of DNA even under native conditions (Fig. 1A, L4). Finally, DNA damage induced by restriction enzyme–like molecules produces multiple bands under native conditions and even smaller fragments under denaturing conditions when combined with damage induced by alkylation or DNase-like molecules (Fig. 1A, L5).

We confirmed that plasmid DNA was stable under diverse anaerobic cultivation conditions including incubation in Gifu medium, which supports the growth of all isolates in our collections (fig. S1). To minimize the damage caused by bacterial DNases that are often produced in the stationary phase of bacterial growth, we measured growth curves for all 122 isolates in our collection and established a time point of exponential phase (T_E) and time point of stationary phase (T_S) (table S1). The isolates were then clustered into seven groups that exhibited similar growth dynamics (fig. S2). We selected two culture conditions for the initial screening: anaerobic co-incubation with DNA to T_S, or anaerobic co-incubation with DNA to T_E followed by aerobic co-incubation to T_S to approximate the oxygen stress encountered in an inflammatory gut environment. Finally, we purified the linearized pUC19 DNA from the bacterial cultures through column purification and performed gel electrophoresis under native and gradient denaturing conditions (0%, 0.2%, 0.4%, and 1% NaOH) (Fig. 1A and fig. S3, A to G).

We used the relative intensity reduction [(RIR), percent] of DNA after co-incubation with bacteria as a general measure of bacterially induced DNA damage (Fig. 1B and table S2). We found that diverse gut microbes exhibited DNA-damaging activities, which suggests that microbiota-mediated genotoxicity may be more widespread than previously acknowledged. Although previously described microbiota-derived genotoxins discovered in a case-by-case manner are primarily produced by gram-negative bacteria (e.g., *E. coli*, *B. fragilis*, *and Klebsiella oxytoca*) (*25*, *26*), we observed that multiple gram-positive microbes from the phyla *Actinobacteria* and *Firmicutes* also caused substantial DNA damage. DNA-damaging activity was largely independent of culture conditions, although select microbes displayed varied genotoxicity in the presence versus absence of oxygen stress (Fig. 1B and table S2).

We selected 24 isolates that exhibited strong DNA-damaging activities in the primary screen and 18 phylogenetically related nongenotoxic isolates for evaluation in a secondary screening (Fig. 1C). We reestablished precise growth curves for each isolate (table S1) and rescreened all 42 isolates under four distinct culture conditions, including co-incubation of DNA with bacterial supernatants collected from anaerobic cultures at T_S (fig. S3, H to J, and table S2). Eighteen isolates that caused DNA damage in our primary screening also exhibited strong genotoxicity upon secondary screening (Fig. 1C and fig. S3, H to J). Moreover, supernatants from most of the selected genotoxic isolates also caused considerable RIR (percent) that were often comparable to co-incubation with live bacteria (Fig. 1C and fig. S3, H to J). Unlike the interstrand cross-links induced by a colibactin-producing *E. coli* strain (K-12 BW25113 containing the *clb* locus, designated as *clb*+ *E. coli*), these 18 isolates exhibited distinct DNA damage patterns after co-incubation with linearized pUC19 DNA. By contrast, a non–colibactin-producing *E. coli* strain (K-12 BW25113 with an empty bacterial artificial chromosome, designated as *clb*− *E. coli*) or medium alone did not induce substantial DNA damage (Fig. 1D and fig. S3K). Notably, most of the newly identified genotoxic isolates induced alkylation or DNase-like DNA damage patterns; we also observed evidence of DNA damage under native gel electrophoresis.

Gut microbes from patients with IBD produce small-molecule metabolites that can induce DNA damage

To determine whether the 18 putative genotoxic bacterial isolates we identified through electrophoresis-based screening (Fig. 1 and fig. S4A) produce genotoxic small molecules that cause DNA damage in human cells, we separated their supernatants (SUP) into small- and large-molecular weight fractions (<3 kDa SUP and >3 kDa SUP, respectively), and evaluated the genotoxicity of these fractions using HeLa cells. Small-molecule metabolites from most selected isolates, including multiple strains of *Bifidobacterium adolescentis* (three isolates), *C. perfringens* (four isolates), *C. ramosum* (phylogenetically related to *C. perfringens*), and *M. morganii* (two isolates), induced increased γ-H2AX, a marker of DNA double-strand breaks (DSBs) (*27*) (fig. S4B). As previously reported, supernatants from *clb*+ *E. coli* failed to induce γ-H2AX, which instead required live bacterial infection (*11*) (fig. S4B). Whereas both small and large molecules from *B. adolescentis* and *C. perfringens* induced increased γ-H2AX, only small-molecule metabolites from *C. ramosum* and *M. morganii* exhibited genotoxicity (fig. S4, B and C). Small-molecule metabolites from *B. adolescentis* and *B. dentium* induced increased apoptosis and necrosis, whereas small-molecule metabolites from all other isolates had minimal impacts on cell viability as measured by cell size and granularity (fig. S4D), Annexin V, or 7-AAD (fig. S4, E to G). Although large molecules from *C. perfringens* induced substantial cell death, likely a result of the effects of clostridial toxins (*28*), small-molecule metabolites from *C. perfringens* caused DNA damage without triggering substantial cell death (fig. S4, B to G). On the basis of these results, we selected *C. perfringens*, *C. ramosum*, and *M. morganii* for further study and confirmed that small-molecule metabolites from these strains induced γ-H2AX in HeLa cells, albeit at a reduced level compared with the well-known DNA-crosslinking chemical cisplatin (Fig. 2A). Small-molecule metabolites from *C. perfringens*, *C. ramosum*, and *M. morganii* also induced cell cycle arrest in HeLa cells (Fig. 2B), further implicating these taxa as potential genotoxin producers.

To enrich genotoxic small molecules, we performed ethyl acetate extractions using supernatants and found that extracts from *C. perfringens* (NWP4), *C. ramosum* (NWP50), *M. morganii* (NWP135), and *clb*+ *E. coli* cultures nicked circular pUC19 plasmid DNA, whereas extracts from *clb*− *E. coli* or medium alone had negligible impacts on DNA integrity (Fig. 2C). Similarly, ethyl acetate extracts from genotoxic species induced γ-H2AX expression in HeLa, HCT116, and MC38 cells (Fig. 2D and fig. S4H) and caused tailing in an alkaline comet unwinding assay (Fig. 2E).

Recent meta analyses have identified cross-cohort microbial signatures associated with CRC, including enrichment in *Clostridiaceae*, *Erysipelotrichaceae*, and *M. morganii* (*29*, *30*). Notably, both *C. perfringens* and *M. morganii* were also increased in IBD patients (specifically in CD patients) compared with healthy controls in data from the Human Microbiome Project (Fig. 2F), suggesting potential roles in IBD patients who are at an increased risk of CRC diagnosis (*21*, *22*). Overall, *M. morganii* is enriched in fecal samples from both IBD and CRC patients and the genotoxicity of this species is restricted to its small-molecule fractions. Therefore, we prioritized *M. morganii* for further genotoxin identification and characterization.

M. morganii *produces genotoxins that are distinct from colibactin*

The biosynthetic machinery involved in the production of microbial metabolites, including previously characterized small-molecule genotoxins, is often encoded by biosynthetic gene clusters (BGCs) (*31*). For example, colibactin production is encoded by a multimodular PKS-NRPS pathway in *E. coli* (*11*, *32*) and tilimycin and tilivalline are encoded by an NRPS pathway in *K. oxytoca* (*26*). However, BGC analyses of genotoxic *C. perfringens*, C. *ramosum*, and *M. morganii* using antiSMASH (*33*) failed to detect any known genotoxin-encoding BGCs (fig. S5A and table S3). While *M. morganii* harbors one NRPS/PKS gene cluster, this BGC is entirely distinct from the *clb* genomic island and *M. morganii* lacks key genes involved in colibactin synthesis, such as *clbI* and *clbP* (fig. S5B) (*34*, *35*). This is consistent with recent

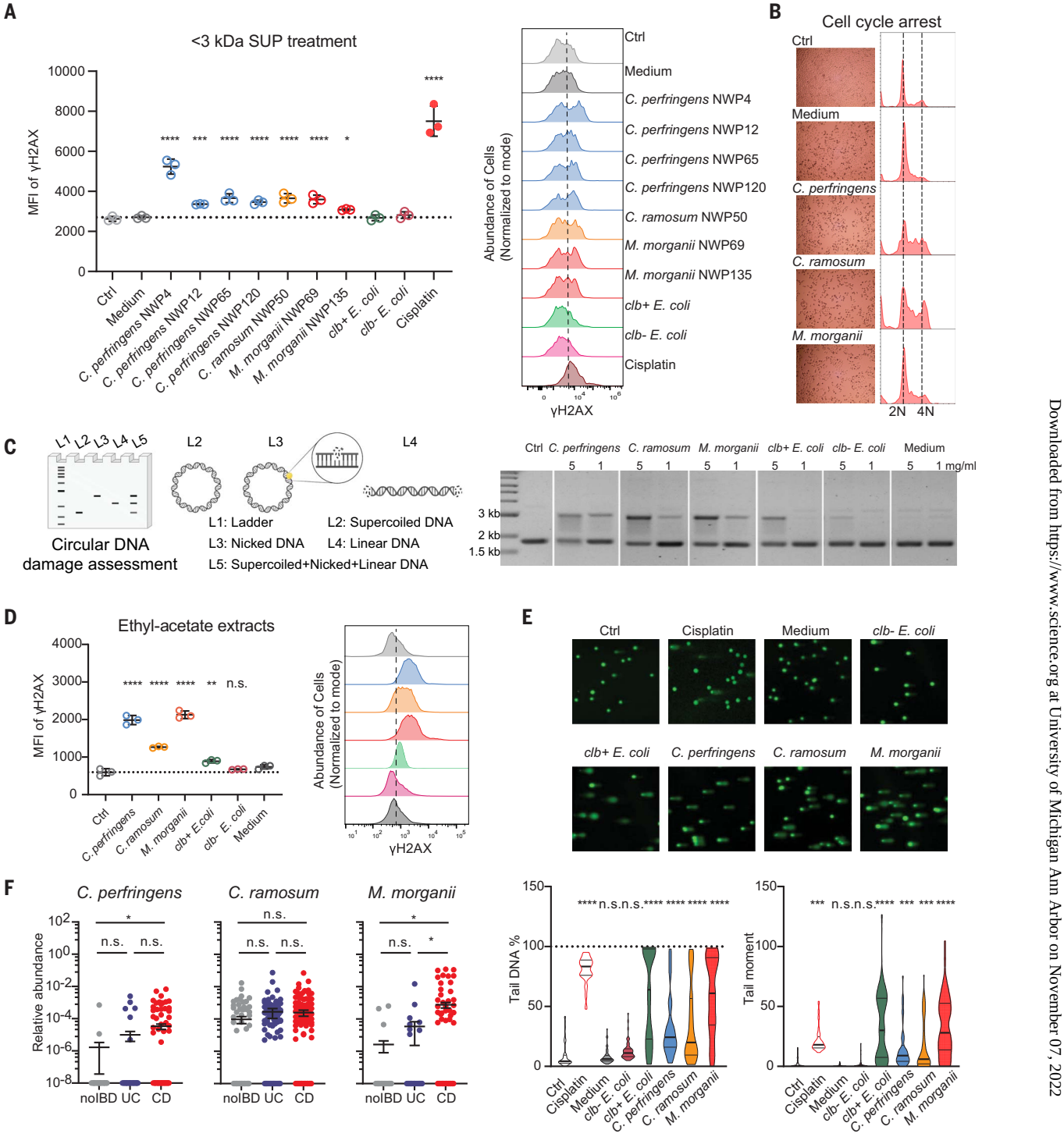


Fig. 2. Small-molecule metabolites produced by gut microbes induce DNA damage. (**A**) Geometric mean fluorescence intensity (MFI) and representative histograms of γ-H2AX staining of HeLa cells treated with 40% (v/v) PBS (ctrl), <3 kDa SUP (small-molecule supernatants) of medium, *C. perfringens*, *C. ramosum*, *M. morganii*, *clb+ E. coli*, *clb− E. coli* isolates, or cisplatin for 5 to 6 hours (n = 3 replicates, N = 3 independent experiments). *P < 0.05; ***P < 0.001; ****P < 0.0001, one-way analysis of variance (ANOVA). (**B**) Representative data of cell cycle arrest evaluated by propidium iodide (PI) staining (n = 3, N = 2). HeLa cells were treated with 40% (v/v) PBS or <3 kDa SUP of medium, *C. perfringens*, *C. ramosum*, or *M. morganii* isolates. (**C**) Assessment of circular pUC19 DNA damage after co-incubation with ethyl-acetate extracts of *C. perfringens*, *C. ramosum*, *M. morganii*, *clb+ E. coli*, *clb− E. coli* supernatants, or medium for 5 to 6 hours (N = 2). Ctrl, control pUC19 DNA in TE buffer. (**D**) MFI of γ-H2AX staining of HeLa cells treated with PBS (ctrl), 5 mg/ml bacterial or medium extracts for 5 to 6 hours (n = 3, N = 2). n.s., not significant; **P < 0.01; ****P < 0.0001, one-way ANOVA. (**E**) Comet assay for genomic DNA damage evaluation. Single-cell gel electrophoresis was performed after treating HeLa cells with PBS (ctrl), cisplatin, bacterial, or medium extracts for 5 to 6 hours (n = 49, N = 1). ***P < 0.001; ****P < 0.0001, one-way ANOVA. (**F**) Relative abundance of *C. perfringens*, *C. ramosum*, and *M. morganii* in data from the Human Microbiome Project. noIBD, healthy controls (n = 429); UC, ulcerative colitis patients (n = 459); CD, Crohn's disease patients (n = 750). *P < 0.05, one-way ANOVA. Data are means ± SEM.

analyses of 69 publicly available *Morganella* genomes, which suggest that *clb* genes are absent in *Morganella* genomes (*36*). The genotoxicity caused by *M. morganii* is also distinct from that caused by colibactin—whereas *clb*+ *E. coli* caused DNA crosslinking, DNA exposed to *M. morganii* displayed a smearing pattern under both native and denaturing conditions (fig. S5C). Finally, as previously reported, colibactin-induced γ-H2AX required live bacterial infection, and supernatants from *clb*+ *E. coli* failed to induce substantial increases in γ-H2AX in cell lines, likely as a result of documented colibactin instability (*11*, *37*). By contrast, *M. morganii* supernatants and small-molecule metabolites elicited substantial increases in γ-H2AX (Fig. 2 and fig. S4). Together, these data suggest that *M. morganii* produces previously undescribed genotoxins that are distinct from colibactin and are readily diffusible.

Isolation and identification of a family of genotoxins derived from M. morganii

To identify specific genotoxins produced by *M. morganii*, we employed a combination of ultraperformance liquid chromatography quadrupole time-of-flight mass spectrometry (UPLC-QTOF-MS)-based untargeted metabolomics, and bioactivity-guided fractionation using small-scale cultures, followed by large-scale cultivation and isolation for unambiguous structure elucidation and genotoxicity analyses (Fig. 3A). We generated an initial candidate ion list of the most abundant *M. morganii*-derived metabolites relative to Gifu medium control (~100 ion features; table S4) through comparative metabolomics. We then performed two rounds of activity-guided fractionation using preparative high-performance liquid chromatography (HPLC) and a circular pUC19 plasmid-based genotoxicity assay, then profiled the resulting fractions and subfractions using UPLC-QTOF-MS-based metabolomics (fig. S6, A and B). To identify potential genotoxins, we excluded ions present in inactive fractions from the initial ion list and ultimately identified four ion features (I to IV) as potential genotoxic hits (Fig. 3B and table S4). To enable structural elucidation and genotoxic activity assessment for these compounds, we performed large-scale cultivation (18 liters) and ethyl acetate extraction of *M. morganii* supernatant based on previously observed retention times that imply relatively low polarity of the compounds of interest. The crude extract was subjected to two rounds of HPLC to generate four semi-pure fractions enriched in the four target ion features (F1 to F4 enriched in I to IV, respectively). One of these fractions (F2) exhibited dose-dependent genotoxicity in a circular pUC19 plasmid-based genotoxicity assay (Fig. 3C). Based on UPLC-QTOF-MS analyses, F2 was primarily comprised of two metabolites with mass/charge ratio (*m/z*) values of 215.1543 (compound **1**, target ion feature II) and 234.1852 (compound **2**, a metabolite that was absent from the initial ion list but was co-enriched during extraction and fractionation) at a ratio of 4:6 (fig. S6C). This fraction was recalcitrant to further purification by preparative HPLC using diverse combinations of stationary and mobile phases. Thus, the chemical structures of the two components were characterized as a mixture using one- and two-dimensional nuclear magnetic resonance (NMR) spectroscopy analyses (Fig. 3D and fig. S7). Compound **1** consists of a conjugation of indole-3-aldehyde (IAld) and the primary amine isoamylamine, forming a functional imine group. Thus, we termed this previously undescribed metabolite indolimine-214.

To confirm which compound exerted the observed genotoxicity, we synthesized both indolimine-214 (**1**) and compound **2**. Formation of the reversible imine functional group in compounds such as indolimine-214 (**1**) varies depending on the chemical environment (e.g., pH, temperature, solvent, and solutes). Therefore, we fractionated fresh synthetic material and assessed the purity of **1** in each fraction using ^{1}H NMR for genotoxicity analysis (fig. S6D). Nonetheless, neither synthetic indolimine-214 (**1**) nor compound **2** alone (up to 1 mg/ml) induced DNA damage in the circular pUC19 plasmid-based genotoxicity assay. However, the mixture of indolimine-214 (**1**) and compound **2** elicited dose-dependent DNA damage (fig. S6D), suggesting that the presence of compound **2** may serve as an adjuvant for indolimine-214 (**1**) (target ion feature II) in the cell-free assay. In a more biologically relevant context, we assessed the genotoxicity of pure synthetic compounds in HeLa cells and found that indolimine-214 (**1**) alone—but not compound **2**—triggered increased γ-H2AX in a dose-dependent manner (Fig. 3E) and induced tailing in an alkaline comet unwinding assay (Fig. 3F). Furthermore, the genotoxicity of synthetic indolimine-214 (**1**) correlated directly with its purity (i.e., in relation to its hydrolytic degradation products; fig. S6E).

Using UPLC-QTOF-MS–based quantification, we found that *M. morganii* produces a high level of indolimine-214 (**1**) in vitro (~40 μg/ml; fig. S8A), which is comparable to the concentrations of the synthetic compound that induced genotoxicity in HeLa cells (Fig. 3, E to F). By contrast, indolimine-214 (**1**) was undetectable in supernatants from wild-type (WT) colibactin-producing (*clbP*+) or isogenic *clbP* mutant non–colibactin-producing (*clbP*–) *E. coli* NC101 strains (*38*) (fig. S8A). This is consistent with the observation that *M. morganii*-induced DNA damage was distinct from the damage caused by colibactin-producing *E. coli* (fig. S5). Cecal contents from mice colonized with *M. morganii*—but not those colonized by *clbP*– *E. coli* NC101—also contained high levels of indolimine-214 (**1**) (fig. S8B). Moreover, in the process of quantifying cecal indolimine-214 (**1**), we observed the presence of two additional new indolimines with similar structures in *M. morganii*-colonized mice: conjugates of IAld with either isobutylamine (indolimine-200, compound **3**, *m/z* 201.1386) or phenethylamine (indolimine-248, compound **4**, *m/z* 249.1386) (Fig. 3G, fig. S8C, and fig. S9). These additional indolimines were also detected within in vitro *M. morganii* bacterial cultures, but not *clbP*– *E. coli* NC101 cultures, as confirmed by synthetic standards (fig. S8D). Finally, synthetic indolimine-200 (**3**) and indolimine-248 (**4**) also triggered increased γ-H2AX in HeLa cells in a dose-dependent manner (Fig. 3H). Taken together, these data show that *M. morganii* produces a family of genotoxic indolimines both in vitro and in vivo.

A previously uncharacterized bacterial decarboxylase is necessary for indolimine synthesis

All *M. morganii*-derived indolimines contain a functional imine group, which is likely derived from the spontaneous condensation of a primary amine (isoamylamine, isobutylamine, and phenethylamine) and the aldehyde of indole-3-aldehyde (IAld) (Fig. 4A). Primary amines are microbial metabolites that can be synthesized from amino acids through a one-step reaction mediated by bacterial decarboxylases (*39*). Based on whole-genome sequencing, *M. morganii* NWP135 encoded 18 predicted decarboxylases, including three decarboxylases (Peg1085, Peg1320, and Peg3098) that were partially homologous to a previously characterized valine decarboxylase from *Streptomyces viridifaciens* (*39*) (Fig. 4B and table S5). We transformed codon-optimized DNA sequences of these three candidate decarboxylases into *E. coli*, induced protein expression with IPTG, and fed each culture with IAld and the relevant amino acid precursors (leucine, valine, or phenylalanine). Peg1085—annotated as pyridoxal-dependent decarboxylase or aspartate aminotransferase (AAT) superfamily (fold type I) in NCBI (Fig. 4B and table S5)—expression in *E. coli* enabled robust production of indolimine-214 (**1**), indolimine-200 (**3**), and indolimine-248 (**4**) based on QTOF-MS identification (Fig. 4C). Therefore, the *aat* gene encoding AAT_I (Peg1085) enables indolimine synthesis.

To evaluate whether the *aat* gene is necessary for indolimine synthesis in *M. morganii*, we constructed a random mutagenesis library of *M. morganii* NWP135 using EZ-Tn5 transposomes and isolated an isogenic *aat* mutant (*40*) (Fig. 5A). Briefly, after optimizing transposome electrotransformation, we picked ~16,000 colonies and mapped the transposon insertion sites by combining combinatorial

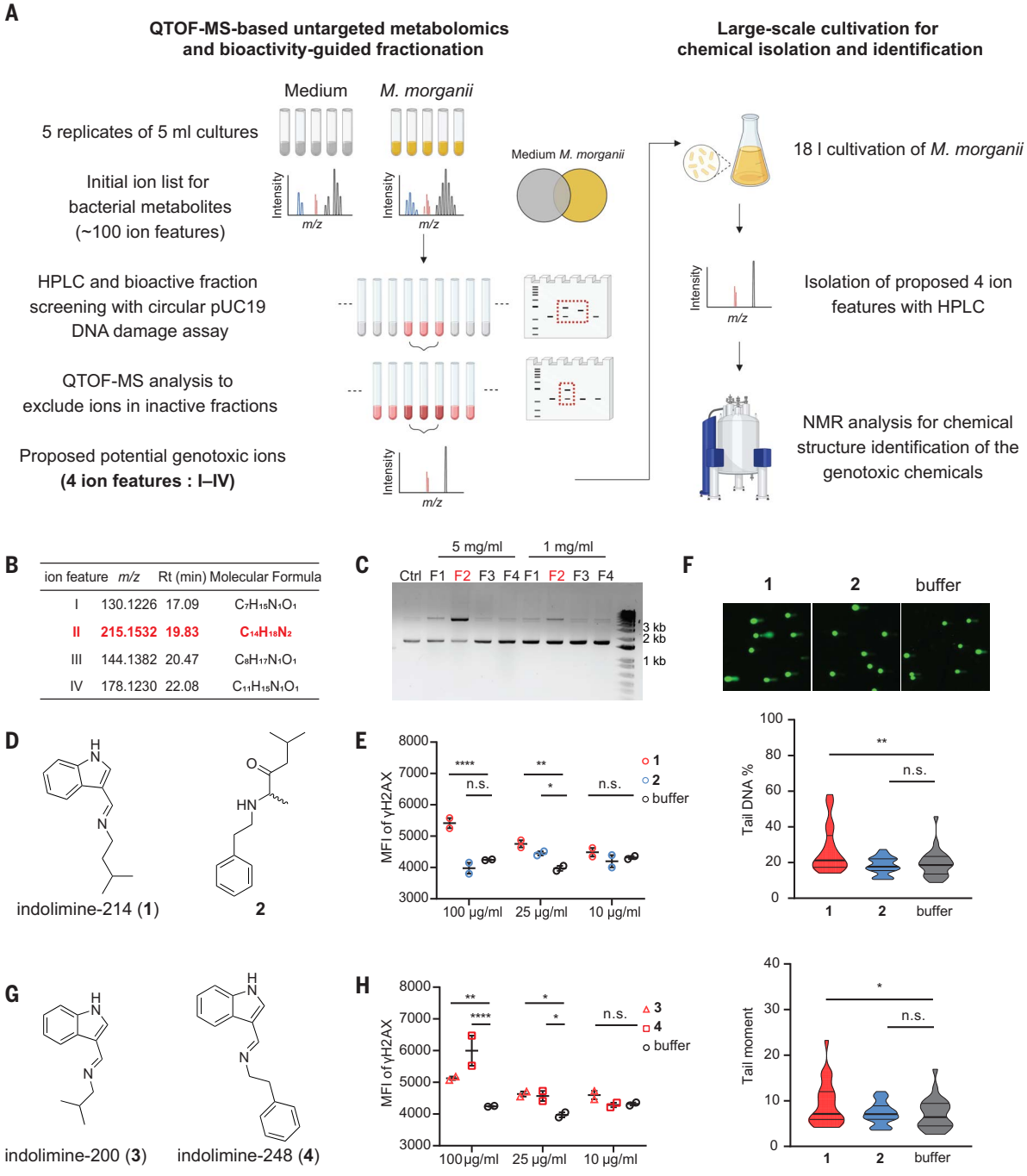
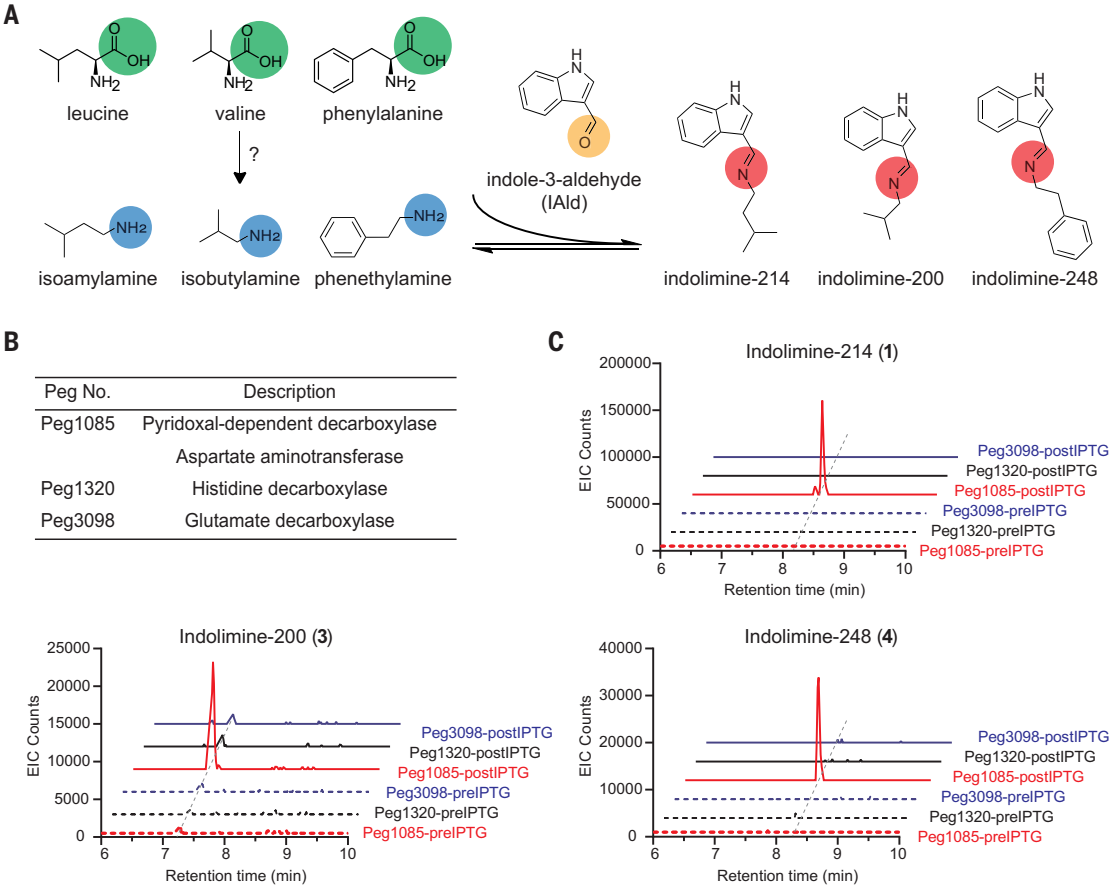


Fig. 3. Isolation and identification of a family of genotoxic metabolites derived from *M. morganii*. (**A**) Overview of isolation and identification of genotoxins derived from *M. morganii*. (**B**) Four proposed candidate ion features initially detected from *M. morganii* bacterial cultures. Rt, retention time. (**C**) Assessment of circular pUC19 DNA damage after co-incubation overnight with F1 to F4 fractions enriched with ion features I to IV, respectively ($N = 1$). Ctrl, control pUC19 DNA in TE buffer. (**D**) Chemical structures of compounds indolimine-214 (**1**) and **2**. (**E**) MFI of γ-H2AX staining of HeLa cells treated with synthetic compounds at indicated concentrations for 5 hours ($n = 2$, $N = 3$). n.s., not significant; *$P < 0.05$; **$P < 0.01$; ****$P < 0.0001$, two-way ANOVA. (**F**) Single-cell genomic DNA comets in HeLa cells after treatment with 100 μg/ml synthetic compounds for 5 to 6 hours ($n = 25$, $N = 1$). *$P < 0.05$; **$P < 0.01$, one-way ANOVA. (**G**) Chemical structures of compounds indolimine-200 (**3**) and indolimine-248 (**4**). (**H**) MFI of γ-H2AX staining of HeLa cells treated with synthetic compounds at indicated concentrations for 5 hours ($n = 2$, $N = 2$). *$P < 0.05$; **$P < 0.01$; ****$P < 0.0001$, two-way ANOVA. Data are means ± SEM.

pooling and transposon sequencing (Tn-Seq) using custom primers adapted from Knockout Sudoku (fig. S10A and table S6) (*41*, *42*). We identified one mutant strain (*aat*– *M. morganii*) with a transposon insertion 7 base pairs (bp) after the *aat* start codon (*aat*–; Fig. 5B and fig. S10B). As compared with WT (*aat*+) *M. morganii*, *aat*– *M. morganii* failed to produce indolimines (Fig. 5C) despite exhibiting normal growth dynamics (Fig. 5D). *aat*– *M. morganii* also failed to induce DNA damage in a cell-free linearized plasmid DNA electrophoresis assay (Fig. 5E) or cell-based γ-H2AX assay (Fig. 5F) as compared with *aat*+ *M. morganii*. Therefore, the *aat* gene is essential



Peg No.	Description
Peg1085	Pyridoxal-dependent decarboxylase Aspartate aminotransferase
Peg1320	Histidine decarboxylase
Peg3098	Glutamate decarboxylase

Fig. 4. *M. morganii* decarboxylase encoded by the *aat* gene enables indolimine synthesis. (**A**) Proposed biosynthesis of indolimines in *M. morganii*. (**B**) Candidate proteins in *M. morganii* NWP135 with significant orthology to valine decarboxylase. Peg, protein encoding gene. (**C**) QTOF-MS identification of indolimine-214 (**1**), indolimine-200 (**3**), and indolimine-248 (**4**) in *E. coli* BL21(DE3) ($N = 2$). *E. coli* cells were separately transformed with the plasmid pET28-Peg harboring codon-optimized DNA sequences of Peg1085, Peg1320, or Peg3098. Indolimines were detected after IPTG induction and feeding with precursors (IAld and leucine, valine and phenylalanine, respectively). preIPTG, bacterial supernatants before IPTG induction; post-IPTG, bacterial supernatants after IPTG induction and precursor feeding.

for indolimine synthesis and *M. morganii*-induced genotoxicity.

Indolimine-producing M. morganii *induces increased gut permeability and exacerbates colon tumor burden in gnotobiotic mice*

To evaluate the impacts of *M. morganii*-derived indolimines in vivo, we first compared the effects of *aat*+ and *aat*– *M. morganii* strains on the intestinal epithelium using monocolonized mice. We found that mice colonized with *aat*+ *M. morganii* and fed IAld and leucine (precursors of indolimine-214) exhibited significantly increased intestinal permeability as compared with *aat*– *M. morganii*–colonized mice (Fig. 6A). Furthermore, RNA-seq of colonic intestinal epithelial cells and gene ontology analyses of differentially expressed genes revealed an increased expression of genes involved in cell cycle regulation, chromosome segregation, and DNA biosynthesis in mice colonized with *aat*+ *M. morganii* (Fig. 6B, fig. S10, C and D, and table S7). Together, these data suggest that the indolimines may cause abnormal DNA replication and IEC proliferation in vivo.

We next evaluated whether indolimine-producing *aat*+ and non–indolimine-producing *aat*– *M. morganii* strains induced differential inflammatory responses, with WT colibactin-producing *clbP*+ and non–colibactin-producing *clbP*– *E. coli* NC101 strains (*38*) as controls (fig. S11A). We found that all groups of monocolonized mice exhibited similar levels of bacterial colonization and roughly equivalent inflammatory responses after treatment with dextran sulfate sodium (DSS) (fig. S11, B to F), indicating that genotoxin production did not significantly alter bacterial colonization or gross inflammatory phenotypes in a model of acute colitis.

To test the potential effects of indolimine production on colon tumorigenesis, we colonized gnotobiotic mice with *aat*+ *M. morganii* or *aat*– *M. morganii* in the context of a mock community of human gut microbes, followed by treatment with azoxymethane (AOM) and three cycles of DSS (Fig. 6C). We selected seven nongenotoxic human gut isolates based on our prior in vitro cell-free genotoxicity screening (Fig. 1 and fig. S11G) to construct a mock community (Geno- community) and used WT colibactin-producing *clbP*+ or its isogenic mutant non–colibactin-producing *clbP*– *E. coli* NC101 strains (*38*) as genotoxin-producing or non–genotoxin-producing positive and negative controls (Fig. 6C). As expected, colonization with *clbP*+ *E. coli* NC101 induced increased colorectal tumor burden as compared with *clbP*– *E. coli* NC101; similarly, mice colonized with indolimine-producing *aat*+ *M. morganii* also exhibited an increased number of tumors and tumor scores (indicating overall tumor burden) as compared with mice colonized with non–indolimine-producing *aat*– *M. morganii* (Fig. 6D and fig. S11H). Moreover, mice colonized with *aat*+ *M. morganii* exhibited an increased ratio of adenomatous lesions with high-grade dysplasia (Fig. 6E). Nonetheless, mice colonized with *aat*+ or *aat*– *M. morganii* exhibited similar levels of *M. morganii* colonization (fig. S11I) and intestinal inflammation, as measured by colon length (fig. S11J), levels of fecal lipocalin 2 (fig. S11K) and histopathology (fig. S11L). These data suggest that genotoxic indolimine-producing *aat*+ *M. morganii* exacerbates colon tumorigenesis in this model, but does not elicit substantial

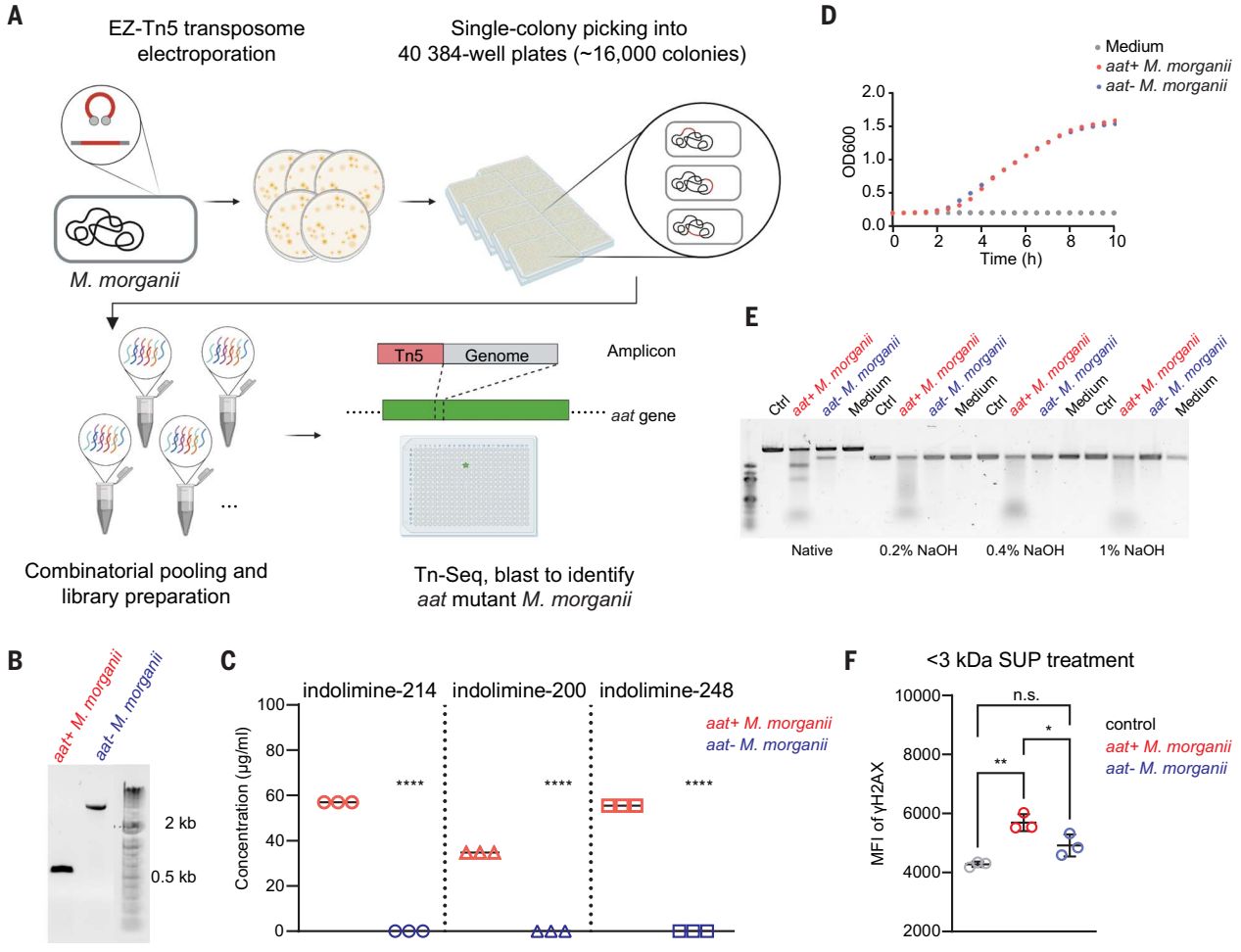


Fig. 5. An isogenic *M. morganii aat* mutant fails to produce indolimines and lacks genotoxicity in vitro. (**A**) Schematic pipeline of random mutagenesis library construction and mutant identification. (**B**) Gel result of PCR products of *aat* gene in *aat*− or *aat*+ *M. morganii* ($N = 2$). (**C**) QTOF-MS quantification of indolimine-214 (**1**), indolimine-200 (**3**), and indolimine-248 (**4**) in bacterial supernatants of *aat*+ *M. morganii* or *aat*− *M. morganii* ($n = 3$, $N = 2$). ****$P < 0.0001$, Student's t test. (**D**) Growth curves of *aat*+ *M. morganii* or *aat*− *M. morganii* ($n = 3$, $N = 1$). (**E**) Gel electrophoresis of cell-free DNA damage assay ($N = 2$). Linearized pUC19 DNA was co-incubated with medium, *aat*+ *M. morganii*, or *aat*− *M. morganii* for 7 to 8 hours, isolated through column purification and treated with or without NaOH (0%, 0.2%, 0.4%, and 1%) before evaluating DNA integrity through gel electrophoresis. (**F**) MFI of γ-H2AX in HeLa cells treated with 40% (v/v) <3 kDa SUP of *aat*+ *M. morganii*, or *aat*− *M. morganii* for 5 to 6 hours ($n = 3$, $N = 2$). n.s., not significant; *$P < 0.05$; **$P < 0.01$, one-way ANOVA. Data are means ± SEM.

increases in intestinal inflammation as compared with a nongenotoxic control.

Finally, we mined the Cancer Microbiome database (*43*) and found that *Morganella* exhibited increased prevalence or abundance in primary gastrointestinal (GI) tumors, including colon adenocarcinomas (TCGA-COAD), rectum adenocarcinomas (TCGA-READ), and stomach adenocarcinomas (TCGA-STAD), as compared with multiple non-GI tumors (fig. S12A). *Morganella* was also enriched in tumors from TCGA-READ and TCGA-STAD as compared with adjacent solid normal tissues (fig. S12B), consistent with a recent report that *Morganella* is enriched in cancerous tissues as compared with luminal contents (*44*). Notably, we also found that *aat* was conserved across nearly all *M. morganii* strains with full genome sequences in the NCBI database (51 of 52 genomes; fig. S12C), suggesting that indolimine production may be a conserved feature of *M. morganii*. Overall, our data—in combination with prior reports—suggest that genotoxic indolimines from *M. morganii* may serve as tumor-inducing factors in humans.

Discussion

Aside from a small number of case studies (*26*, *45*), the taxonomic distribution and repertoire of small-molecule genotoxins produced by microbiota remain mostly unexplored. We undertook a systematic evaluation of the genotoxicity of a diverse selection of human gut microbes based on prior evidence that colibactin-producing *E. coli* induces DNA damage and facilitates intestinal tumorigenesis (*14*). Through our investigation, we (i) found that diverse taxa from the human gut microbiota exhibited genotoxicity; (ii) identified and characterized a previously undescribed family of genotoxic *M. morganii*-derived small molecules termed the indolimines; (iii) decoded the indolimine production pathway in *M. morganii* and generated an isogenic non–indolimine-producing *M. morganii* mutant; and (iv) found that indolimine-producing *M. morganii* exacerbated colon tumorigenesis in gnotobiotic mice.

By revealing the existence of a previously uncharted universe of microbiota-derived genotoxins and defining the indolimines as a previously undescribed family of bioactive microbiota-derived small molecules, these studies imply an expanded role for genotoxic metabolites in CRC. We focused most of our studies on *M. morganii* as it is enriched in both CRC (*29*, *30*, *43*, *44*) and IBD patients, the latter of which are at increased risk of CRC diagnosis (*21*, *22*). However, additional genotoxic species identified in our initial screens may also contribute to CRC. Indeed, genotoxic *C. perfringens*

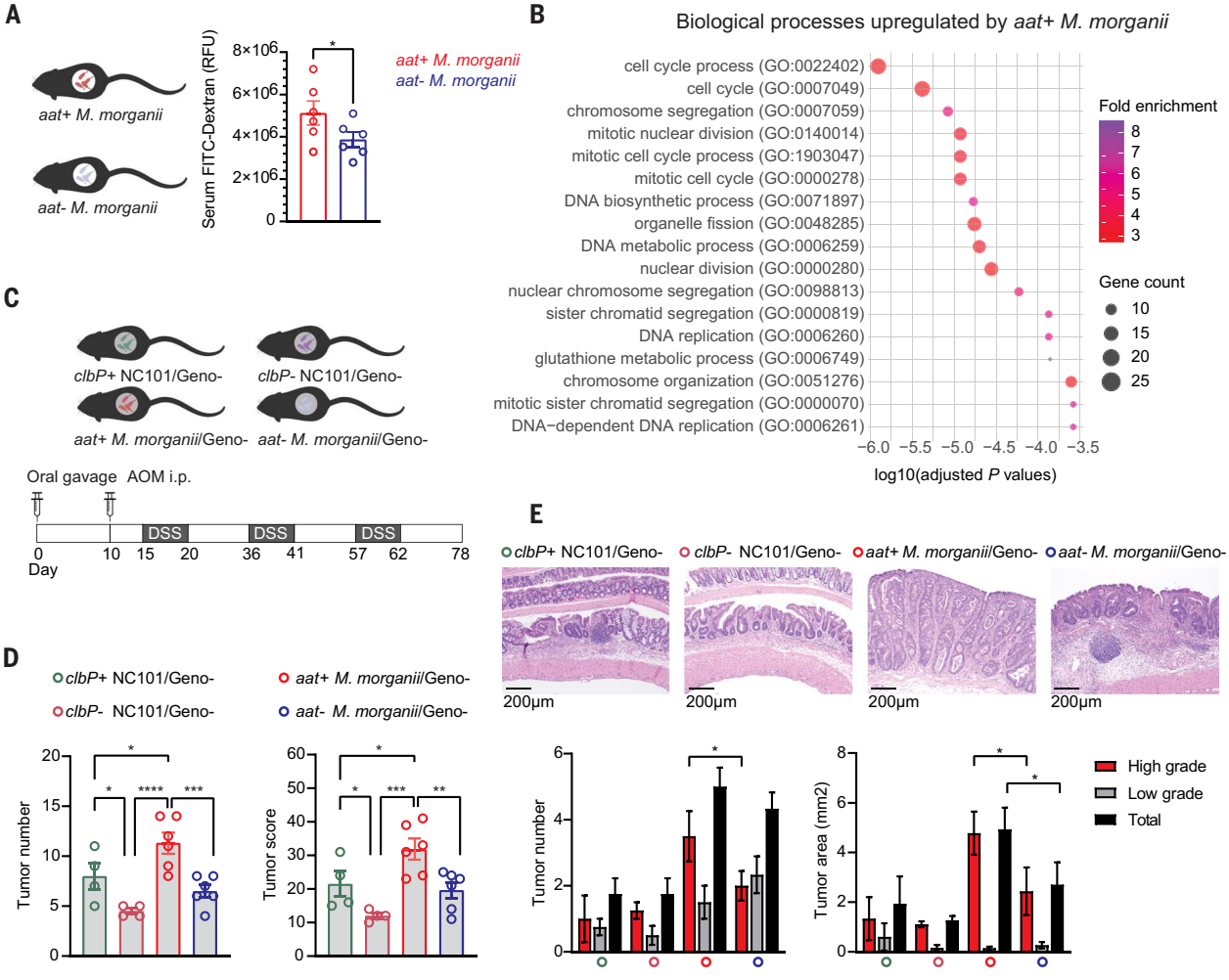


Fig. 6. Indolimine-producing *M. morganii* induces increased gut permeability and exacerbates colon tumor burden in gnotobiotic mice. (**A**) Evaluation of intestinal permeability of mice colonized with *aat+* or *aat−* *M. morganii* based on serum FITC-Dextran RFU (relative fluorescence units) (n = 6). *P < 0.05, Student's t-test. (**B**) GO-Slim biological process analysis based on the gene ontology overrepresentation test for colonic epithelial cells from mice colonized with *aat+* (n = 5) or *aat−* *M. morganii* (n = 6). Data are pooled from N = 1 experiment. (**C**) Schematic of experimental design for AOM/DSS model in age-matched gnotobiotic mice colonized with *clbP+ E. coli* NC101 (n = 4), *clbP− E. coli* NC101 (n = 4), *aat+ M. morganii* (n = 6) or *aat− M. morganii* (n = 6) with nongenotoxic mock community (Geno-, constructed with 7 nongenotoxic isolates) (N = 2). (**D** and **E**) Tumor number and tumor score (D), representative tumor histology images (scale bar = 200 μm) and tumor grade evaluation [(E) tumor number and tumor area per section of tissues], in gnotobiotic mice colonized with *clbP+ E. coli* NC101, *clbP− E. coli* NC101, *aat+ M. morganii* or *aat− M. morganii* with Geno- community. n.s., not significant; *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001, one-way ANOVA. Data are means ± SEM.

and *C. ramosum* strains also promoted colorectal tumor burden in gnotobiotic mice as compared with a nongenotoxic mock community (fig. S12, D to G), but did not produce indolimines (fig. S12H), suggesting that additional microbiota-derived genotoxins remain to be characterized.

Notably, somatic mutations can be detected in human colonic epithelial cells even in early life, which suggests persistent mutagenesis throughout the lifespan of an individual (*46*, *47*), and colitis-related expansions of mutated clones may influence both IBD pathogenesis and CRC susceptibility (*48*). Furthermore, although CRC patients display increased carriage of *clb+ E. coli*, *clb+* taxa (including *E. coli* relatives such as *Klebsiella* species) are also found in healthy individuals (*49*). Recent studies also revealed that increased epithelial oxygenation during colitis could drive *clb+ E. coli* expansion through aerobic respiration, increasing colibactin-mediated CRC-inducing activity (*50*). These observations support a model whereby genotoxic gut microbes contribute to CRC development by persistently inducing DNA damage in host epithelial cells, which synergizes with chronic inflammation in the gut microenvironment, along with additional environmental factors, and eventually facilitates the initiation and progression of CRC.

Microbiota-derived genotoxins may also affect diverse aspects of host biology beyond tumor initiation. Recent studies revealed that colibactin also influences gut microbiome composition (*51*), exacerbates lymphopenia and septicemia (*52*), triggers prophage induction through the bacterial SOS response (*53*), and restricts *Vibrio cholerae* colonization (*54*). Thus, commensal-derived genotoxins, including the indolimines, may also mediate diverse biological functions. Overall, our studies underscore the power of function-based assessments of the microbiome to provide new insights into the diverse impacts of indigenous microbes on host biology and disease susceptibility.

REFERENCES AND NOTES

1. N. Keum, E. Giovannucci, Global burden of colorectal cancer: Emerging trends, risk factors and prevention strategies. *Nat. Rev. Gastroenterol. Hepatol.* **16**, 713–732 (2019). doi: 10.1038/s41575-019-0189-8; pmid: 31455888
2. K. W. Jasperson, T. M. Tuohy, D. W. Neklason, R. W. Burt, Hereditary and familial colon cancer. *Gastroenterology* **138**,

2044–2058 (2010). doi: 10.1053/j.gastro.2010.01.054; pmid: 20420945
3. C. A. Brennan, W. S. Garrett, Gut Microbiota, Inflammation, and Colorectal Cancer. *Annu. Rev. Microbiol.* **70**, 395–411 (2016). doi: 10.1146/annurev-micro-102215-095513; pmid: 27607555
4. H. Tilg, T. E. Adolph, R. R. Gerner, A. R. Moschen, The Intestinal Microbiota in Colorectal Cancer. *Cancer Cell* **33**, 954–964 (2018). doi: 10.1016/j.ccell.2018.03.004; pmid: 29657127
5. W. S. Garrett, Cancer and the microbiota. *Science* **348**, 80–86 (2015). doi: 10.1126/science.aaa4972; pmid: 25838377
6. A. Koh, F. De Vadder, P. Kovatcheva-Datchary, F. Bäckhed, From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites. *Cell* **165**, 1332–1345 (2016). doi: 10.1016/j.cell.2016.05.041; pmid: 27259147
7. M. R. Rubinstein *et al*., Fusobacterium nucleatum promotes colorectal carcinogenesis by modulating E-cadherin/β-catenin signaling via its FadA adhesin. *Cell Host Microbe* **14**, 195–206 (2013). doi: 10.1016/j.chom.2013.07.012; pmid: 23954158
8. J. Allen, C. L. Sears, Impact of the gut microbiome on the genome and epigenome of colon epithelial cells: Contributions to colorectal cancer development. *Genome Med.* **11**, 11 (2019). doi: 10.1186/s13073-019-0621-2; pmid: 30803449
9. A. C. Goodwin *et al*., Polyamine catabolism contributes to enterotoxigenic *Bacteroides fragilis*-induced colon tumorigenesis. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15354–15359 (2011). doi: 10.1073/pnas.1010203108; pmid: 21876161
10. E. Buc *et al*., High prevalence of mucosa-associated *E. coli* producing cyclomodulin and genotoxin in colon cancer. *PLOS ONE* **8**, e56964 (2013). doi: 10.1371/journal.pone.0056964; pmid: 23457644
11. J.-P. Nougayrède *et al*., *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. *Science* **313**, 848–851 (2006). doi: 10.1126/science.1127059; pmid: 16902142
12. J. C. Arthur *et al*., Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science* **338**, 120–123 (2012). doi: 10.1126/science.1224820; pmid: 22903521
13. A. Cougnoux *et al*., Bacterial genotoxin colibactin promotes colon tumour growth by inducing a senescence-associated secretory phenotype. *Gut* **63**, 1932–1942 (2014). doi: 10.1136/gutjnl-2013-305257; pmid: 24658599
14. G. Cuevas-Ramos *et al*., *Escherichia coli* induces DNA damage in vivo and triggers genomic instability in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11537–11542 (2010). doi: 10.1073/pnas.1001261107; pmid: 20534522
15. M. Xue *et al*., Structure elucidation of colibactin and its DNA cross-links. *Science* **365**, eaax2685 (2019). doi: 10.1126/science.aax2685; pmid: 31395743
16. M. R. Wilson *et al*., The human gut bacterial genotoxin colibactin alkylates DNA. *Science* **363**, eaar7785 (2019). doi: 10.1126/science.aar7785; pmid: 30765538
17. C. Pleguezuelos-Manzano *et al*., Mutational signature in colorectal cancer caused by genotoxic pks$^+$ E. coli. *Nature* **580**, 269–273 (2020). doi: 10.1038/s41586-020-2080-8; pmid: 32106218
18. P. J. Dziubańska-Kusibab *et al*., Colibactin DNA-damage signature indicates mutational impact in colorectal cancer. *Nat. Med.* **26**, 1063–1069 (2020). doi: 10.1038/s41591-020-0908-2; pmid: 32483361
19. R. R. da Silva, P. C. Dorrestein, R. A. Quinn, Illuminating the dark matter in metabolomics. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 12549–12550 (2015). doi: 10.1073/pnas.1516878112; pmid: 26430243
20. N. W. Palm *et al*., Immunoglobulin A coating identifies colitogenic bacteria in inflammatory bowel disease. *Cell* **158**, 1000–1010 (2014). doi: 10.1016/j.cell.2014.08.006; pmid: 25171403
21. O. Olén *et al*., Colorectal cancer in ulcerative colitis: A Scandinavian population-based cohort study. *Lancet* **395**, 123–131 (2020). doi: 10.1016/S0140-6736(19)32545-0; pmid: 31929014
22. O. Olén *et al*., Colorectal cancer in Crohn's disease: A Scandinavian population-based cohort study. *Lancet Gastroenterol. Hepatol.* **5**, 475–484 (2020). doi: 10.1016/S2468-1253(20)30005-4; pmid: 32066530
23. N. Bossuet-Greif *et al*., The Colibactin Genotoxin Generates DNA Interstrand Cross-Links in Infected Cells. *mBio* **9**, 02393-17 (2018). doi: 10.1128/mBio.02393-17; pmid: 29559578
24. K. S. Gates, An overview of chemical processes that damage cellular DNA: Spontaneous hydrolysis, alkylation, and reactions with radicals. *Chem. Res. Toxicol.* **22**, 1747–1760 (2009). doi: 10.1021/tx900242k; pmid: 19757819
25. C. M. Dejea *et al*., Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria. *Science* **359**, 592–597 (2018). doi: 10.1126/science.aah3648; pmid: 29420293
26. K. Unterhauser *et al*., *Klebsiella oxytoca* enterotoxins tilimycin and tilivalline have distinct host DNA-damaging and microtubule-stabilizing activities. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 3774–3783 (2019). doi: 10.1073/pnas.1819154116; pmid: 30808763
27. L. J. Kuo, L. X. Yang, Gamma-H2AX - a novel biomarker for DNA double-strand breaks. *In Vivo* **22**, 305–309 (2008). pmid: 18610740
28. J. C. Freedman, A. Shrestha, B. A. McClane, Clostridium perfringens Enterotoxin: Action, Genetics, and Translational Applications. *Toxins* **8**, 73 (2016). doi: 10.3390/toxins8030073; pmid: 26999202
29. A. M. Thomas *et al*., Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. *Nat. Med.* **25**, 667–678 (2019). doi: 10.1038/s41591-019-0405-7; pmid: 30936548
30. J. Wirbel *et al*., Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. *Nat. Med.* **25**, 679–689 (2019). doi: 10.1038/s41591-019-0406-6; pmid: 30936547
31. E. E. Shine, J. M. Crawford, Molecules from the Microbiome. *Annu. Rev. Biochem.* **90**, 789–815 (2021). doi: 10.1146/annurev-biochem-080320-115307; pmid: 33770448
32. G. Schneditz *et al*., Enterotoxicity of a nonribosomal peptide causes antibiotic-associated colitis. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 13181–13186 (2014). doi: 10.1073/pnas.1403274111; pmid: 25157164
33. K. Blin *et al*., antiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* **47** (W1), W81–W87 (2019). doi: 10.1093/nar/gkz310; pmid: 31032519
34. E. P. Trautman, A. R. Healy, E. E. Shine, S. B. Herzon, J. M. Crawford, Domain-Targeted Metabolomics Delineates the Heterocycle Assembly Steps of Colibactin Biosynthesis. *J. Am. Chem. Soc.* **139**, 4195–4201 (2017). doi: 10.1021/jacs.7b00659; pmid: 28240912
35. L. Zha *et al*., Colibactin assembly line enzymes use S-adenosylmethionine to build a cyclopropane ring. *Nat. Chem. Biol.* **13**, 1063–1065 (2017). doi: 10.1038/nchembio.2448; pmid: 28805802
36. H. Wami *et al*., Insights into evolution and coexistence of the colibactin- and yersiniabactin secondary metabolite determinants in enterobacterial populations. *Microb. Genom.* **7**, 000577 (2021). doi: 10.1099/mgen.0.000577; pmid: 34128785
37. E. E. Shine *et al*., Model Colibactins Exhibit Human Cell Genotoxicity in the Absence of Host Bacteria. *ACS Chem. Biol.* **13**, 3286–3293 (2018). doi: 10.1021/acschembio.8b00714; pmid: 30403848
38. S. Tomkovich *et al*., Locoregional Effects of Microbiota in a Preclinical Model of Colon Carcinogenesis. *Cancer Res.* **77**, 2620–2632 (2017). doi: 10.1158/0008-5472.CAN-16-3472; pmid: 28416491
39. D. I. Kim, T. U. Chae, H. U. Kim, W. D. Jang, S. Y. Lee, Microbial production of multiple short-chain primary amines via retrobiosynthesis. *Nat. Commun.* **12**, 173 (2021). doi: 10.1038/s41467-020-20423-6; pmid: 33420084
40. Y. Veeranagouda, F. Husain, H. M. Wexler, Transposon mutagenesis of the anaerobic commensal, Bacteroides fragilis, using the EZ:TN5 transposome. *FEMS Microbiol. Lett.* **333**, 94–100 (2012). doi: 10.1111/j.1574-6968.2012.02602.x; pmid: 22639975
41. I. A. Anzai, L. Shaket, O. Adesina, M. Baym, B. Barstow, Rapid curation of gene disruption collections using Knockout Sudoku. *Nat. Protoc.* **12**, 2110–2137 (2017). doi: 10.1038/nprot.2017.073; pmid: 28906493
42. P. J. Lariviere *et al*., An Essential Regulator of Bacterial Division Links FtsZ to Cell Wall Synthase Activation. *Curr. Biol.* **29**, 1460–1470 (2019). doi: 10.1016/j.cub.2019.03.066; pmid: 31031115
43. G. D. Poore *et al*., Microbiome analyses of blood and tissues suggest cancer diagnostic approach. *Nature* **579**, 567–574 (2020). doi: 10.1038/s41586-020-2095-1; pmid: 32214244
44. W. Chen, F. Liu, Z. Ling, X. Tong, C. Xiang, Human intestinal lumen and mucosa-associated microbiota in patients with colorectal cancer. *PLOS ONE* **7**, e39743 (2012). doi: 10.1371/journal.pone.0039743; pmid: 22761885
45. M. W. Dougherty, C. Jobin, Shining a Light on Colibactin Biology. *Toxins* **13**, 346 (2021). doi: 10.3390/toxins13050346; pmid: 34065799
46. I. Martincorena, P. J. Campbell, Somatic mutation in cancer and normal cells. *Science* **349**, 1483–1489 (2015). doi: 10.1126/science.aab4082; pmid: 26404825
47. H. Lee-Six *et al*., The landscape of somatic mutation in normal colorectal epithelial cells. *Nature* **574**, 532–537 (2019). doi: 10.1038/s41586-019-1672-7; pmid: 31645730
48. S. Olafsson *et al*., Somatic Evolution in Non-neoplastic IBD-Affected Colon. *Cell* **182**, 672–684.e11 (2020). doi: 10.1016/j.cell.2020.06.036; pmid: 32697969
49. J. Putze *et al*., Genetic structure and distribution of the colibactin genomic island among members of the family Enterobacteriaceae. *Infect. Immun.* **77**, 4696–4703 (2009). doi: 10.1128/IAI.00522-09; pmid: 19720753
50. S. A. Cevallos *et al*., Increased Epithelial Oxygenation Links Colitis to an Expansion of Tumorigenic Bacteria. *mBio* **10**, 02244-19 (2019). doi: 10.1128/mBio.02244-19; pmid: 31575772
51. S. Tronnet *et al*., The Genotoxin Colibactin Shapes Gut Microbiota in Mice. *MSphere* **5**, 00589-20 (2020). doi: 10.1128/mSphere.00589-20; pmid: 32611705
52. I. Marcq *et al*., The genotoxin colibactin exacerbates lymphopenia and decreases survival rate in mice infected with septicemic *Escherichia coli*. *J. Infect. Dis.* **210**, 285–294 (2014). doi: 10.1093/infdis/jiu071; pmid: 24489107
53. J. E. Silpe, J. W. H. Wong, S. V. Owen, M. Baym, E. P. Balskus, The bacterial toxin colibactin triggers prophage induction. *Nature* **603**, 315–320 (2022). doi: 10.1038/s41586-022-04444-3; pmid: 35197633
54. J. Chen *et al*., A commensal-encoded genotoxin drives restriction of *Vibrio cholerae* colonization and host gut microbiome remodeling. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2121180119 (2022). doi: 10.1073/pnas.2121180119; pmid: 35254905

ACKNOWLEDGMENTS

We thank A. L. Goodman, D. Bajic, B. Lim, Y. Surovtseva, and S. Umlauf, as well as the Yale Microbial Sciences Institute and Yale Center for Molecular Discovery for assistance with assembly of the transposon mutant libraries, the Yale Center for Genome Analysis for sequencing services, and the Yale Animal Resource Center for veterinary services. We also thank B. M. Barstow, P. Chien, and B. Aldikacti for technical advice on Knockout Sudoku and Tn-Seq. **Funding:** This work was supported by a Yale Cancer Center Team Challenge award funded by the National Cancer Institute of the National Institutes of Health under award 3P30CA016359 (to N.W.P., S.B.H., and J.M.C.) as well as R01CA215553 to S.B.H. and J.M.C. Approximately 35% of the funding for this research project (~$200,000) was financed with NIH funds; the remainder was financed by nongovernmental sources. N.W.P. also gratefully acknowledges support from the Common Fund of the National Institutes of Health (DP2DK125119), Leona M. and Harry B. Helmsley Charitable Trust (3083), Chan Zuckerberg Initiative, Michael J. Fox Foundation for Parkinson's Research, Ludwig Family, Mathers Foundation, Pew Charitable Trust, NIA, and NIGMS (R01AG068863 and RM1GM141649), and F. Hoffmann-La Roche Ltd. The funders of this work had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Yale Center for Molecular Discovery is supported in part by an NCI Cancer Center Support grant NIH P30 CA016359. Equipment and libraries were supported in part by the Program in Innovative Therapeutics for Connecticut's Health. **Author contributions:** Y.C. and N.W.P. designed the study and wrote the manuscript with input from all authors. Y.C., J.O., and M.X. designed and performed experiments and analysis. W.J.H. performed histopathological scoring. J.W., J.A.G.-H., T.A.R., A.L.M., and D.S. helped with experiments or contributed essential reagents. S.B.H., J.M.C., and N.W.P. participated in study design and supervised research. **Competing interests:** N.W.P. is a co-founder of Artizan Biosciences, Inc. and Design Pharmaceuticals, Inc. N.W.P. has received research funding for unrelated studies from Artizan Biosciences, Inc. and F. Hoffmann-La Roche AG. Y.C., J.O., J.M.C., M.X., S.B.H., and N.W.P. are inventors on a provisional patent application submitted by Yale University that covers the use of microbiome-targeted therapeutics to treat or prevent commensal genotoxin-induced disease. **Data and materials availability:** Bulk RNA-seq data, fecal 16S rRNA profiling of microbiota for AOM/DSS tumor experiments, and trimmed Tn-seq data for the *M. morganii* mutant library are available at NCBI Bioproject: PRJNA879774. Related analysis methods are available in the supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm3233
Materials and Methods
Figs. S1 to S44
Tables S1 to S8
References (*55*–*62*)
MDAR Reproducibility Checklist

View/request a protocol for this paper from Bio-protocol.

Submitted 9 September 2021; resubmitted 10 July 2022
Accepted 19 September 2022
10.1126/science.abm3233

Science

Commensal microbiota from patients with inflammatory bowel disease produce genotoxic metabolites

Yiyun CaoJoonseok OhMengzhao XueWon Jae HuhJiawei WangJaime A. Gonzalez-HernandezTyler A. RiceAnjelica L. MartinDeguang SongJason M. CrawfordSeth B. HerzonNoah W. Palm

Science, 378 (6618), eabm3233. • DOI: 10.1126/science.abm3233

A new class of bacterial genotoxins

Individuals with inflammatory bowel disease are at increased risk of developing colorectal cancer compared with the general population. The gut microbiome is among the many factors that can influence tumorigenesis, in part by modulating the immune system and producing microbial metabolites. Cao *et al*. developed a functional screen to test whether gut bacteria from patients with inflammatory bowel disease have genotoxic effects (see the Perspective by Puschhof and Sears). The authors discovered a family of DNA damage–inducing microbial metabolites called indolimines, which were produced by the Gram-negative bacteria *Morganella morganii*. In a mouse model of colon cancer, *M. morganii* exacerbated tumor burden, but a mutant form of the bacteria unable to produce indolimine did not. This diverse series of genotoxic small molecules from the human microbiome may play a role in intestinal tumorigenesis. —PNK

Science (ISSN) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.