

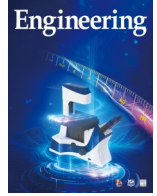


ELSEVIER

Contents lists available at ScienceDirect

Engineering

journal homepage: [www.elsevier.com/locate/eng](http://www.elsevier.com/locate/eng)



Research  
Microecology—Article

## 结直肠癌黏膜组织来源的大肠杆菌菌株间遗传和功能差异研究

常宇骁<sup>a,#</sup>, 李享<sup>a,#</sup>, 丁磊<sup>b,#</sup>, 杨超<sup>a,d,#</sup>, 潘志远<sup>a</sup>, 韩妮<sup>a</sup>, 崔玉军<sup>a</sup>, 智发朝<sup>c</sup>, 杨瑞馥<sup>a</sup>, 高宏<sup>b,\*</sup>, 毕玉晶<sup>a,\*</sup>

<sup>a</sup> State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, Beijing 100071, China;

<sup>b</sup> Beijing Shijitan Hospital, Capital Medical University, Beijing 100038, China;

<sup>c</sup> Guangdong Provincial Key Laboratory of Gastroenterology, Institute of Gastroenterology of Guangdong Province, & Department of Gastroenterology, Nanfang Hospital, Southern Medical University, Guangzhou 510515, China;

<sup>d</sup> Shenzhen Center for Disease Control and Prevention, Shenzhen 518055, China

### ARTICLE INFO

#### Article history:

Received 9 October 2020

Revised 11 February 2021

Accepted 16 March 2021

Available online 30 September 2021

#### 关键词

结直肠癌  
肠道菌群  
培养组学  
大肠杆菌  
菌株水平

### 摘要

结直肠癌是全球第三大癌症。宏基因组学已被广泛用于基于细菌属或种水平的比较来分析肠道微生物群与结直肠癌之间的关系,为结直肠癌发展中的生态失调提供了证据。然而,这种分析并不能为我们提供菌株水平的信息,进而理解一种细菌在结直肠癌发生发展中菌株水平的作用。在本文中,我们利用培养组学方法分离结直肠癌黏膜样本,并选择了158株大肠杆菌,通过系统发育分析和炎症诱导实验,揭示它们在基因组学和功能上的差异。基因组比较可以将这些菌株分为5个系统群。选择代表性菌株进行THP-1细胞(人白血病单核细胞)Transwell实验以及动物实验,结果显示不同菌株刺激后,细胞因子水平有显著性差异。进一步的生物信息学分析揭示了不同系统群间单核苷酸多态性、基因和代谢途径的不同特征,这些结果有助于了解这些菌株之间的表型差异。细菌菌株在基因组学和功能上的差异表明,菌株水平上的功能差异可以进一步了解宿主与肠道细菌的相互作用机制。

© 2021 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. 引言

结直肠癌是最常见的癌症之一,也是全球第三大癌症死亡原因。许多研究已经证实了肠道菌群和结直肠癌发展之间的联系,其中大多数研究是基于生物信息学分析[1–3]。16S rRNA基因的扩增子测序和宏基因组测序是研究肠道微生物组的主要方法,已被广泛用于鉴定可能的结直肠癌致病菌,以及寻找结直肠癌潜在的诊断标志物[4]。当细菌被证明与结直肠癌相关后,就会用细胞或者动物实验来验证这种相关性。通过测序技术来确定微生物的生理

状态和功能是比较困难的;此外大多数研究中用到的菌株通常是从保藏中心购买的,而不是分离自临床患者[5–6]。

同种细菌的菌株之间也存在很大的表型差异。比如,菌株水平的差异会影响膳食化合物的代谢,如半乳糖和不易消化的纤维[7–8]。细菌介导的药物代谢也可能因菌株而异,从而影响药物的疗效和毒性[9]。尽管许多技术已经被用于从数千个与人类相关的细菌宏基因组中提取物种和亚种水平的信息,但使用不依赖于培养的方法在菌株水平上准确区分这些特征是极其困难的[10–11]。培养组学是一种使用多种培养条件来获得细菌,并采用MALDI-

\* Corresponding authors.

E-mail addresses: [gaohongdoctor@sina.com](mailto:gaohongdoctor@sina.com) (H. Gao), [byj7801@sina.com](mailto:byj7801@sina.com) (Y. Bi).

# These authors contributed equally to this work.

2095-8099/© 2021 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

英文原文: *Engineering* 2022, 16(9): 210–219

引用本文: Yuxiao Chang, Xiang Li, Lei Ding, Chao Yang, Zhiyuan Pan, Ni Han, Yujun Cui, Fachao Zhi, Ruifu Yang, Hong Gao, Yujing Bi. Genetic and Functional Differences of *Escherichia coli* Strains from Colorectal Cancer Mucosal Tissues. *Engineering*, <https://doi.org/10.1016/j.eng.2021.03.028>

TOF 质谱和 16S rRNA 测序来鉴定细菌的方法[12]。该技术已成为扩展我们对微生物组理解的关键方法之一，同时也是获得目标菌株，并在菌株水平上研究微生物群与宿主之间相互作用机制的重要方法。如最近的研究显示，Sorbara 等[13]从人类样本中培养了 273 个毛螺科菌株，包括 11 个属和 27 个种，并证明了它们在种间和种内水平上的多样性，证实了微生物群在菌株水平上研究的意义。

在肠道菌群和结直肠癌相关性研究中，样本大多数来自于较容易获得的粪便样本，但是肠黏膜相关细菌可能在微生物群与宿主的相互作用中发挥更重要的作用。研究显示，通过 16S rRNA 扩增子测序证实了黏膜微生物群与结直肠癌更为相关，并发现黏膜中肠道微生物群的多样性和组成与粪便中微生物群的多样性和组成不同[14–15]。最近的一项研究指出，肠黏膜样本可能更适合研究微生物群与宿主之间的关系[16]。为了进一步证明黏膜相关细菌在菌株水平上的差异，我们通过培养组学，从 22 例结直肠癌患者的 44 份肠黏膜标本（包括癌症黏膜和癌旁黏膜）中分离出 158 株大肠杆菌，利用二代测序技术对其进行全基因组测序，并在细胞和动物水平上进行细胞因子检测。然后，进一步分析与毒力相关的基因分布和代谢途径，以寻找功能差异的可能机制，证实了不同菌株的大肠杆菌可表现出较大的基因组变异和功能差异。

## 2. 材料和方法

### 2.1. 样本

纳入研究的 22 例结直肠癌患者来自北京世纪坛医院。样本入选标准：①50~80 岁的结直肠癌患者；②既往未接受过放疗；③符合手术治疗标准。取 2 cm<sup>2</sup> 的结直肠癌组织或癌旁组织，在 2 mL 磷酸盐缓冲液（PBS）中完全研磨。在预定的培养条件下分离培养组织中的菌株。

### 2.2. 培养组学

采用培养组方法对细菌进行分离和鉴定。将黏膜组织置于无菌 PBS 中进行研磨；所有样品均在基础培养条件下培养（有氧或厌氧，37 °C，YCFA 固体培养基）；当平板上出现单克隆时，挑取单克隆并转移至 24 孔板中的液体培养基中进行液体扩增。扩增产物加入甘油进行冻存，同时对扩增产物进行固体培养基三区划线。待三区划线长出单克隆后，用 AutOF MS1000 光谱仪（Autobio，郑州）鉴定细菌种类。如果 MALDI-TOF MS 不能准确鉴定菌落，则通过 16S rRNA 基因测序鉴定分离株。16S rRNA 基因测序由擎科生物科技公司（北京）完成。

### 2.3. 生物信息学分析

#### 2.3.1. 大肠杆菌的分类

大肠杆菌基因组测序由诺禾生物技术公司（北京）完成。序列数据已上传至 GenBank (Bioproject pRJ-NA608078)。我们使用了克隆框架软件[17]。基于 8 个管家基因（总共 4095 个核苷酸）的序列对大肠杆菌进行分析和分类[18]。我们使用 Snippy Pipeline v4.3.8 (<https://github.com/tseemann/snippy>) 进行 SNP 调用，使用 *E. coli* STR K-12 子代 MG1655（登录号：U00096）作为参考基因组。位于重复区域的 SNPs 被排除在系统发育分析之外。用 TRF V4 鉴定重复区域，并通过 BLASTN 进行自排列。使用 TreeBest 软件 (<http://treesoft.sourceforge.net/treebest.shtml>) 将全基因组 SNPs 用于构建邻接树，并使用 GG-Tree 对所有系统发育树进行可视化。

#### 2.3.2. 大肠杆菌基因分析

我们用 SPAdes 进行了重新组装。用 Prokka 对组件进行基因注释，并在 Roary 中使用注释结果（GFF3 文件）来识别泛基因组并生成基因存在/不存在矩阵。我们使用 eggNOG-Mapper 来注释泛基因源的直系同源群集（COG）分类。

### 2.4. Transwell 实验

Transwell 平板是一种嵌套结构，它将孔分为上室和下室。嵌套中间的膜具有不同孔径规格。我们选择 0.4 μm 的孔径以确保细菌产物可以通过膜而细菌不能通过。将 THP-1 细胞以每孔 10<sup>6</sup> 个细胞（10<sup>6</sup> 个细胞·mL<sup>-1</sup>）的接种密度在 12 孔 Transwell 平板（Thermo Fisher Scientific）的下室中接种。然后细胞加入添加终浓度为 100 μg·mL<sup>-1</sup> 的佛波醇 12-肉豆蔻酸酯 13-乙酸酯（PMA）（Thermo Fisher Scientific）的培养基中。在 37 °C、5% CO<sub>2</sub> 的湿润空气中孵育 48 h，以刺激细胞分化为巨噬细胞。随后，将大肠杆菌菌株离心后除去细菌培养基，重悬于细胞培养基中。最后，将不同的大肠杆菌菌株以 MOI=10 加入 Transwell 平板的上室（每孔 10<sup>7</sup> 个）。使用 RPMI 1640 培养基作为阴性对照。在 37 °C、含 5% CO<sub>2</sub> 的湿润空气中孵育 3 h 后，收集细胞培养基。THP-1 细胞（人白血病单核细胞）购自北纳生物技术公司（美国），细胞培养基为添加 10% 牛血清（Cleson Scientific，中国）的 RPMI 1640（Thermo Fisher Scientific）。

### 2.5. 动物实验

雌性 BALB/c 小鼠（8 周龄）购自维通利华生物公司

(北京)。将小鼠随机分为5组, 每组5只小鼠: A1、A2、D1、D2和B1。小鼠适应环境一周后, 在其饮水中加入广谱抗生素:  $1\text{ g}\cdot\text{L}^{-1}$ 氨苄青霉素、 $1\text{ g}\cdot\text{L}^{-1}$ 硫酸新霉素、 $1\text{ g}\cdot\text{L}^{-1}$ 甲硝唑和 $0.5\text{ g}\cdot\text{L}^{-1}$ 万古霉素[19]。喂食后, 每只小鼠立即用 $10^9\text{ cfu}\cdot\text{mL}^{-1}$ 大肠杆菌灌胃, 每周两次, 共4周, 每周尾部采血一次。空白对照组给予等体积无菌PBS。

## 2.6. 细胞因子检测

收集 Transwell 装置中 THP-1 细胞的培养上清液, 并用 Bio-Plex 200® 系统 (Bio-Rad, USA) 进行测定。

## 2.7. 分析和统计

对多个组的细胞因子含量进行比较。首先确认每组数据的正态性和方差齐性。如果数据呈正态分布且显示方差齐性, 则使用方差分析 (ANOVA), 而如果数据不呈正态分布或显示方差异质性, 则使用秩和检验 ( $P < 0.05$ )。采用 Pearson's  $\chi^2$  检验比较两组间致病基因的分布情况。两组总样本数定义为  $n$ , 理论频数定义为  $t$ , 当  $n \geq 40$  且所有  $t \geq 5$  时, 采用 Pearson's  $\chi^2$  检验, 当  $n < 40$  或  $t < 1$  时, 采用 Fisher 确切概率检验 ( $P < 0.05$ )。

# 3. 结果

## 3.1. 结直肠癌黏膜来源的大肠杆菌菌株

黏膜是与肠道疾病最相关的组织, 因此我们使用培养组学方法分离和培养了44个黏膜样本。从结果中可以看出, 大肠杆菌是从结直肠癌和邻近黏膜中分离出的最常见的菌种, 它可以从所有44个样品中分离出来 (附录A中的图S1)。所有大肠杆菌菌株和患者信息如表S1所示。

## 3.2. 大肠杆菌在结直肠癌患者肠道微生物群中的遗传多样性

共生大肠杆菌的基因结构是由多种宿主和环境因素决定的。决定其毒力的因素可能反映了其对共生环境的适应。我们选择了158株大肠杆菌进行全基因组测序。在这些菌株中, 来自同一患者的菌株用于评估宿主内的多样性, 来自不同患者的菌株用于宿主间的比较。利用 Clonal Frame 软件对大肠杆菌进行分类分析[17,20], 菌株被分为5个类群, 即A、D、B1、B2和F (图1)。但是来自癌组织和癌旁组织的大肠杆菌菌株在系统发育树上没有明显区分, 表明来自两种组织的菌群之间不存在基因型差异。

### 3.2.1. 宿主内多样性

早期一项研究表明, 在任何时候, 一个人通常携带一

种占优势的微生物群菌株, 该菌株占分离出的菌落的一半以上, 其他菌株也以不同的水平存在[21]。在本研究中, 我们注意到一些患者携带一个优势系统类群。例如, 在患者18中, 大多数菌株属于种系群B1, 然而, 基于单核苷酸多态性 (SNP) 分析, 它们可以进一步区分为不同的分支, 表明这些菌株在基因组水平上的差异。在患者2中, 仅分离出来自系统类群A的菌株。相反, 在患者3中, 菌株被分配到三个系统类群, 没有优势系统类群。

### 3.2.2. 宿主间多样性

从结果中可以看到, 宿主之间菌株水平的细菌多样性很高, 我们没有检测到任何两个患者有完全相同的菌株。这些结果进一步加深了对细菌菌株功能的理解, 这对于细菌在疾病发展和微生态研究中的精确机制至关重要。随后, 我们采用免疫细胞实验和动物实验来进一步评估不同种系代表株的细胞因子诱导效果。

## 3.3. 大肠杆菌不同菌株对 THP-1 细胞的影响

由于肠道微生物群在正常状态下不与免疫细胞直接接触, 因此肠道细菌对免疫细胞的作用主要是通过其代谢产物介导的。因此, 使用 Transwell 方法来研究不同大肠杆菌菌株对免疫细胞的影响[22]。根据大肠杆菌类群分类, 共选择36个代表性菌株 (来自6名患者) 用于实验。Transwell 平板分成上室和下室, 首先将 THP-1 细胞 (人白血病单核细胞) 加入下室, 然后用 PMA 诱导细胞 48 h 以诱导其分化为巨噬细胞, 最后将相应的大肠杆菌菌株加入上室 (感染复数 [MOI] = 10)。孵育 3 h 后, 收集细胞培养基以测定白细胞介素  $1\beta$  (IL- $1\beta$ )、肿瘤坏死因子  $\alpha$  (TNF- $\alpha$ ) 和白细胞介素 6 (IL-6) 的水平 [图 2 (a)]。

首先, 我们比较了来自不同种类群的大肠杆菌菌株对巨噬细胞分泌这三种细胞因子的影响 [图 2 (b) ~ (d)]。结果表明, 5 个类群之间存在显著差异。D 群菌株刺激后细胞因子的分泌量低于其他群菌株, 而 A 群菌株刺激后细胞因子的分泌量始终高于 D 群菌株。我们还研究了样品来源和样品分类是否影响巨噬细胞分泌细胞因子。结果表明, 即使菌株来自同一患者, 不同种系的菌株之间也存在显著差异; 但即使样本来自不同的患者, 细胞因子诱导效应在同一系统类群中也是相似的 [图 2 (e) ~ (g)]。总之, 菌株的种系对巨噬细胞分泌细胞因子的影响比菌株的来源更明显。

## 3.4. 不同大肠杆菌菌株对小鼠的影响

由于在细胞实验中 A 群和 D 群菌株之间存在显著差异, 我们从系统发育树分支中选择了两株大肠杆菌进行动



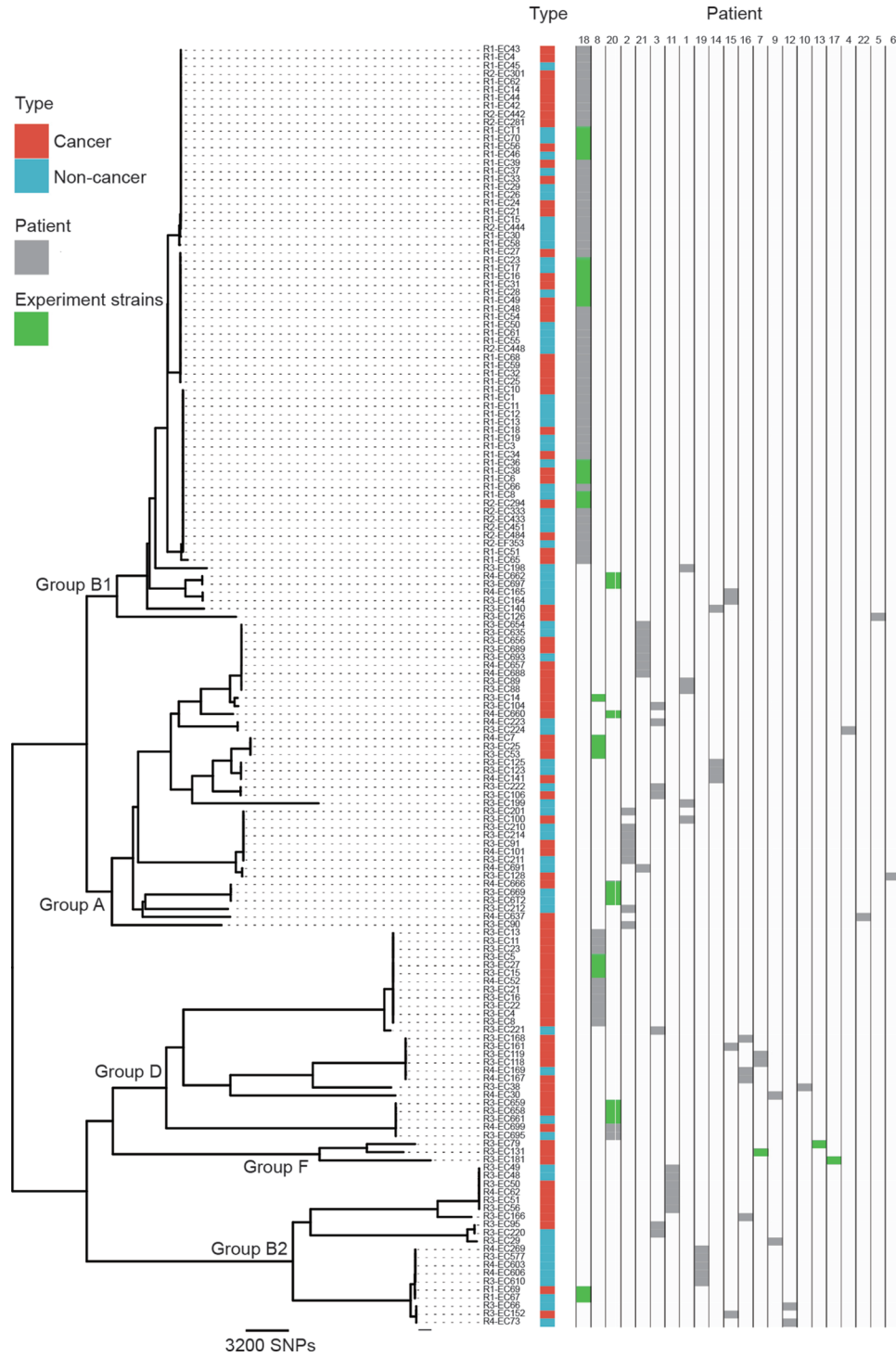


图1. 结直肠癌相关大肠杆菌的系统发育树。对来自22名患者的158株大肠杆菌进行了分析。这些菌株可分为5个主要的系统发育群：A、B1、B2、D和F。红色表示来自癌组织的菌株，蓝色表示来自非癌组织的菌株。对于进一步的Transwell实验菌株用绿色标注。

物实验分析。我们将四个菌株分别命名为A1、A2、D1和D2，以及空白对照组BL。我们首先用联合抗生素喂养小鼠1周以清理其肠道微生物。在抗生素处理后，用相应的大肠杆菌菌株每周灌胃小鼠两次，持续4周，并且在每周第二次灌胃两天后收集血样[图3 (a)]。在实验过程中，各组之间未观察到体重差异（图S2）。实验结束时，与对

照组相比，四个实验组的结直肠病理无明显变化（图S3）。每周检测一次血清中IL-1 $\beta$ 、TNF- $\alpha$ 和IL-6的含量[图3 (b) ~ (d)]。在第1周，细胞因子的变化与细胞实验的结果一致，而在第2周和第3周，4个实验组和对照组之间没有观察到差异。但在第4周，血清中细胞因子的浓度与第1周相反，用D组菌株刺激后的细胞因子浓度

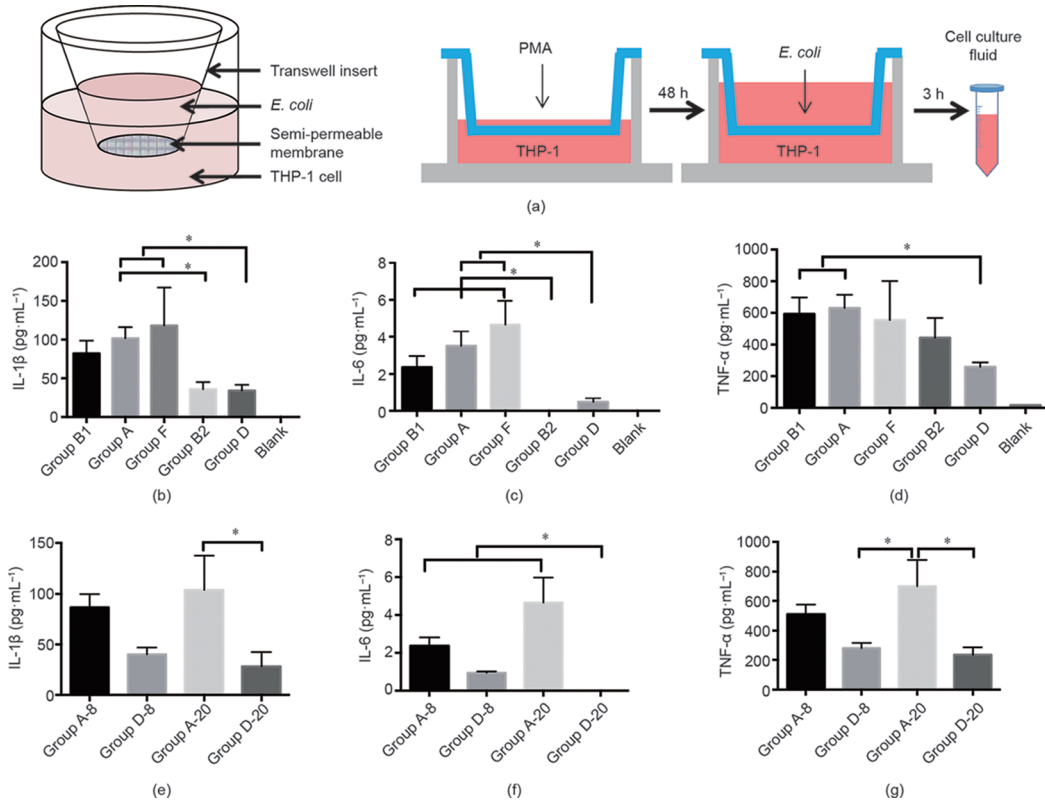


图2. (a) Transwell 实验设计; (b) 各组细胞培养基中 IL-1 $\beta$  含量的比较; (c) 各组细胞培养基中 IL-6 含量的比较; (d) 各组细胞培养基中 TNF- $\alpha$  含量的比较; (e) 细胞培养基中 IL-1 $\beta$  含量的组内比较; (f) 细胞培养基中 IL-6 含量的组内比较; (g) 细胞培养基中 TNF- $\alpha$  含量的组内比较。\* $P \leq 0.05$ 。

显著高于用 A 组菌株刺激后的细胞因子浓度。

### 3.5. 大肠杆菌 A 群和 D 群菌株的遗传分析

为了探索两个大肠杆菌系统群之间对于细胞因子诱导差异的可能机制, 我们进一步分析了细胞实验中使用的 14 个菌株 (6 个来自系统群 D, 8 个来自系统群 A) 的基因组差异, 以了解毒性相关的基因分布和代谢途径。我们在谱系特异性核心基因中发现了 33 248 个 SNPs, 在谱系特异性辅助基因中发现了 348 个 SNPs [图 3 (a) 和表 S2]。当我们将 SNP 相关基因定位到京都基因和基因组百科全书 (KEGG) 通路时, 两个系统类群之间有 37 条显著不同通路 (表 S3)。在 348 个基因中, 278 个基因特异性地存在于系统类群 D 中, 70 个基因特异性地存在于系统类群 A 中。这些基因大部分与细菌代谢有关, 少数是与细胞侵袭和黏附有关的毒力相关基因。与核心基因相关的途径包括氨基酸、核苷酸、碳水化合物和无机离子的转运和代谢。与辅助基因相关的途径仅涉及碳水化合物转运和代谢 (图 S4)。

前期研究显示, 在大肠杆菌中, 有 100 多个基因与毒力有关[23–31]。我们筛选了 103 个毒力相关基因或基因组岛, 并比较了 A 和 D 类群菌株中的这些基因。然后, 我们计算了这些基因在每个系统类群中的比例 (表 S4)。在图 4 (b) ~ (d) 中展示了显著不同的基因, 包括测序菌

株[图 4 (b)]、细胞实验中使用的菌株[图 4 (c)]以及在动物实验中使用的菌株[图 4 (d)]。14 个基因 (*chuA*、*ipaH*、*fmlA*、*fimA*、*fimC*、*ecpA*、*ecpD*、*kpsS*、*ydeR*、*FimH*、*iss*、*aufC*、*lpfA* 和 *stfD*) 在 D 群中的频率高于 A 群 ( $P \leq 0.025$ ), 12 个基因 (*yfcV*、*ChiA*、*elfC*、*stcD*、*papC*、*iucC*、*iucD*、*iutA*、*iha*、*sat*、*papF* 和 *papGII*) 在 A 群菌株中比在 D 群菌株中更普遍 ( $P \leq 0.050$ )。据报道, 系统类群 D 中的几个基因 (如 *chuA*、*ipaH*、*fmlA*) 与黏附和侵袭有关, 而系统类群 A 中的 *yfcV* 与黏附有关。

虽然我们没有发现来自不同患者的任何相同菌株, 但我们在来自不同患者的菌株中观察到相同的毒力基因谱 (图 S5)。例如, 分别来自患者 21 和患者 4 的菌株 R4-EC223 和 R3-EC224 具有相同的已知毒力基因谱。又比如, 菌株 R3-EC201 和 R3-EC100 不仅来自不同的患者, 而且来自不同的黏膜部位 (癌症和邻近组织), 也具有相同的毒力基因谱。

## 4. 讨论

目前, 大多数关于肠道细菌和结直肠癌的报道都采用宏基因组学方法鉴定相关细菌, 然后使用美国典型培养物

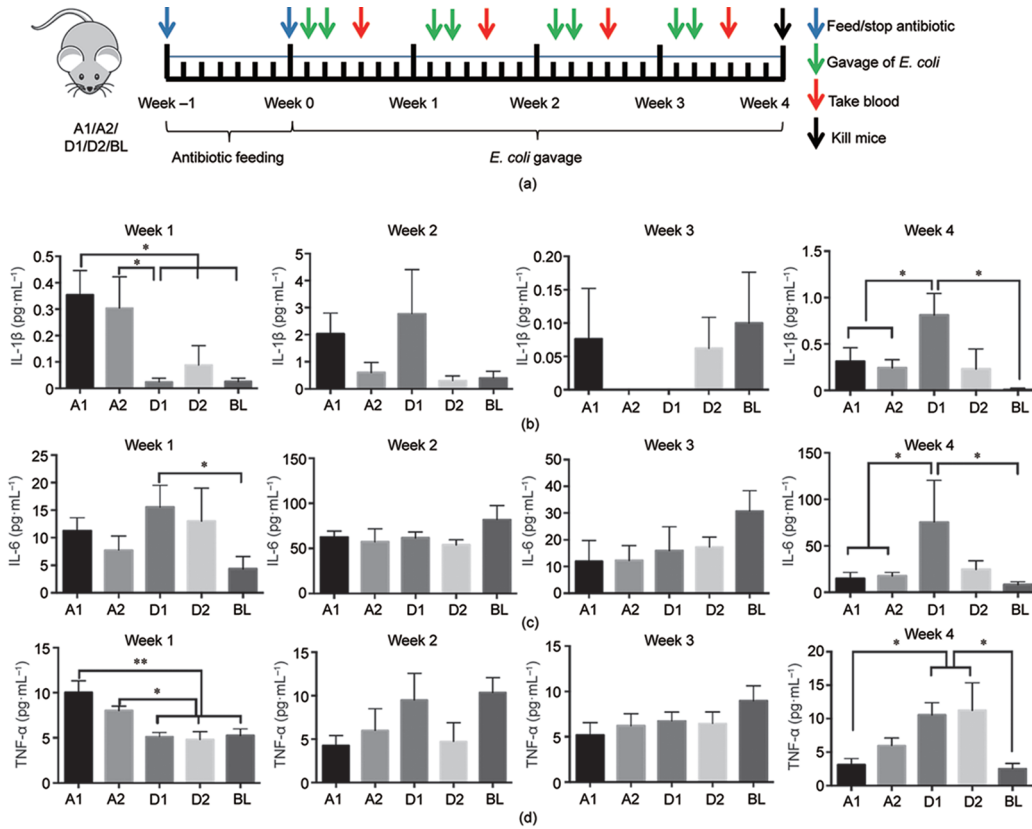


图3. (a) 小鼠实验设计；(b) 灌胃后1~4周小鼠血清中IL-1 $\beta$ 含量的比较；(c) 灌胃后1~4周小鼠血清中IL-6含量的比较；(d) 灌胃后1~4周小鼠血清中TNF- $\alpha$ 含量的比较。\* $P \leq 0.05$ ；\*\* $P \leq 0.01$ 。

保藏中心 (ATCC) 的菌株进行进一步验证。例如，具核梭杆菌 (*Fusobacterium nucleatum*) 被发现与结直肠癌相关，菌株 ATCC 23726 和 25586 被用于在无菌小鼠中诱导肿瘤[5,32]。虽然 AOM 在肠黏膜损伤后成功诱导了肿瘤，但使用结直肠癌患者的原始菌株进行验证实验，这可能更有助于理解靶细菌在癌症发展中的致病作用。

在本研究中，分离自结直肠癌患者黏膜组织的大肠杆菌菌株，从基因组和功能上均体现了菌株差异。大肠杆菌是人体肠道中常见的一种细菌，有些会引起胃肠道疾病[33]，也会引起泌尿系统[34]、神经系统感染[35]。同时有宏基因组数据也表明克罗恩病患者个体中存在多种大肠杆菌菌株[36]。根据基因组分类结果，158株大肠杆菌被分为A、D、B1、B2和F共5个类群。我们发现，来自不同患者的大肠杆菌菌株属于不同的类群，表明不同宿主间，肠道微生物群中的细菌菌株具有高度多样性。随后，我们进一步找出了这些类群的功能差异。

细胞因子是反映肿瘤患者免疫反应的重要指标。IL-1 $\beta$ 、TNF- $\alpha$  和 IL-6 是在癌症中表达变化显著的三种细胞因子，它们的水平甚至可能与肿瘤发展的阶段相关[37–38]。因此，使用 Transwell 实验，我们首先检测了来自不

同类群的不同大肠杆菌菌株对巨噬细胞的刺激情况。结果显示，不同类群间差异明显，A群和D群诱导的细胞因子水平差异最为显著。如果来自同一患者的菌株属于不同的系统群，这些菌株在诱导免疫细胞分泌细胞因子方面是否存在差异？我们比较来自患者8和患者20的A群和D群中的菌株。结果显示来自同一种类群的大肠杆菌菌株对细胞因子的诱导没有差异。但是，大肠杆菌不同类群的菌株刺激免疫细胞后，其分泌的细胞因子水平明显不同（例如，来自同一患者的A-20和D-20诱导的IL-1 $\beta$ 和TNF- $\alpha$ 分泌有显著性差异）。这些结果提示，对于肠道微生物群的研究要在菌株水平展开，以准确了解细菌菌株在肠道中的作用。

接下来，我们用来自A群和D群的菌株刺激小鼠，来观察是否可以在动物模型中检测到类似的结果。在细菌刺激之前，所有的小鼠都接受了抗生素治疗，以清除它们的肠道微生物群[19]。菌株灌胃一周后，A、D两组小鼠血清IL-1 $\beta$ 和TNF- $\alpha$ 水平有显著差异，与细胞实验结果一致。在接下来的两周内，用不同菌株刺激的小鼠，其血清细胞因子分泌没有差异。然而，在灌胃后的第4周，与A组相比，D组的菌株再次在小鼠中引起显著的细胞因子分泌。这是一个有趣的表型结果，我们试图在基因水平上对

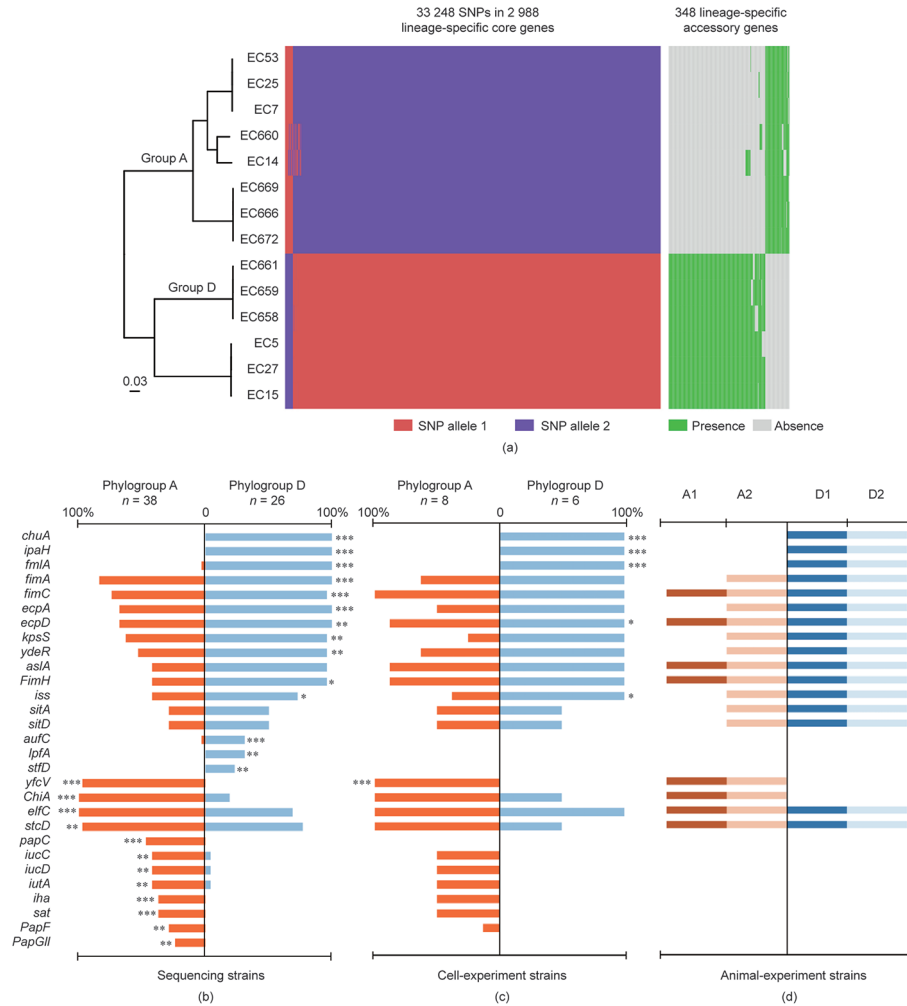


图4. (a) 谱系特异性核心基因和谱系特异性辅助基因的单核苷酸多态性 (SNPs) 在A群和D群之间的比较。(b) ~ (d) 致病基因在A群和D群中的分布: (b) 所有测序菌株; (c) 仅用于细胞实验的菌株; (d) 仅用于小鼠实验的菌株。数字表示含有基因菌株数量相对于该类群的菌株总数的百分比。图中展示的是有统计学显著差异的基因。\* $P \leq 0.05$ ; \*\* $P \leq 0.01$ ; \*\*\* $P \leq 0.001$ 。

其进行解释。我们首先分析了A群和D群中存在和缺失的基因。总共鉴定了352个类群特异性辅助基因，这些基因主要存在于D群中。在A群中，大多数大肠杆菌菌株是产肠毒素性大肠杆菌 (ETEC)，其通过表面蛋白和可能的其他表面结构附着于宿主上皮而引起疾病[23]。然而，黏附侵袭性大肠杆菌 (AIEC) 与炎症性肠病有关，可产生引起肠道炎症和病变的毒素[27]，主要分布在类群D和B2中。研究表明，表达*yfcV*基因的大肠杆菌比不表达*yfcV*基因的大肠杆菌更容易黏附到黏膜上，从而引起炎症。*yfcV*是一种黏附基因，编码假定的伴侣蛋白的主要亚基[39]。*yfcV*基因在A组中的表达显著高于D组，这可能与A组在细胞实验中和刺激小鼠后第一周血清中的细胞因子表达显著升高有关。然而，随着感染时间的延长，具有比A群更多毒力基因的D群可能诱导宿主产生更多的细胞因子。早期研究也提到，AIECs会有多种毒力因子来促进其对细胞的侵袭[40]。

除了基因缺失，SNPs的差异也会在菌株水平上影响细菌的功能。例如，与黏附相关的S层糖蛋白的*gmhB*基因在系统发育D群中表达被中断[41]。又比如，A群中的*yghJ*基因（也称为*SSIE*）框内缺失破坏，可能导致其失活。研究表明，*yghJ*基因与细菌分泌系统和生物膜形成有关，触发NF- $\kappa$ B和MAP激酶信号通路，从而诱导炎症因子的表达[42–43]。因此，A群和D群中的SNPs也可能导致表型差异。

有研究表明，不同的细菌菌株对人类健康的影响不同。目前人类肠道中大多数细菌物种的活体分离菌株及基因组信息仍然有限，特别是在菌株水平的研究更为有限。此外，大多数用于菌株水平研究的菌株来自不同的个体，而我们对来自同一人体的菌株多样性的了解非常有限。一些研究指出，同一物种的菌株之间存在功能差异，这可能会影响人类健康[44–45]。因此，使用培养法从结肠直肠癌患者的肿瘤黏膜及其邻近组织中分离出黏膜相关细菌，这

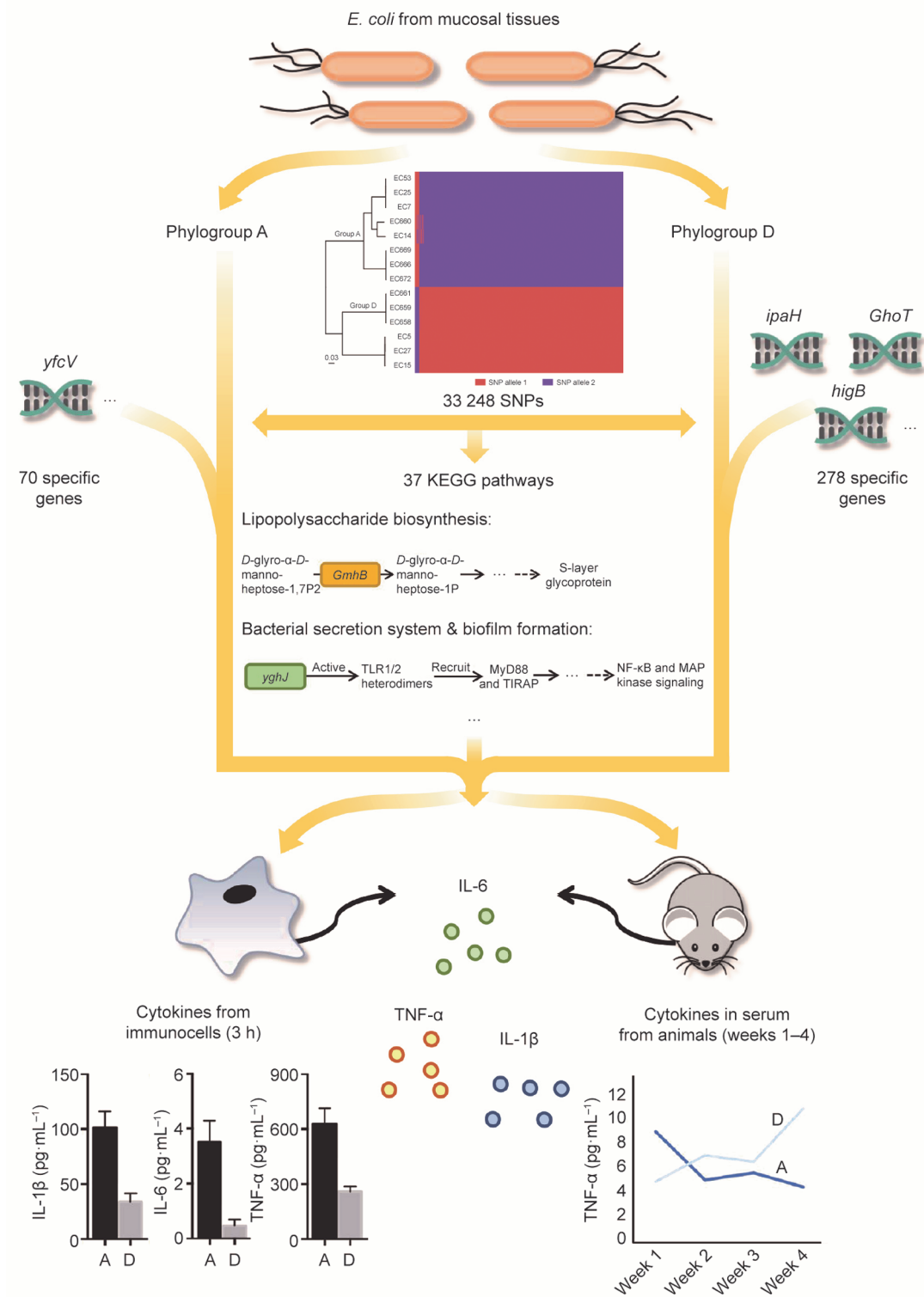


图5. 菌株水平大肠杆菌对宿主的影响及机制。

可以为未来的功能研究提供了结直肠癌相关的肠道微生物菌库。本研究中对大肠杆菌菌株进行分离培养、测序，及分析它们对细胞和宿主的功能，进而发现了可能的作用机制（图5），为在菌株水平上进行深入的筛选、分析和验证提供了研究基础。

## 致谢

本研究得到了国家自然科学基金重点项目（81790632）、国家自然科学基金（31970863）、广州市创新领军团队计划（201809010014）的支持。



## 作者贡献

杨瑞馥构思了本实验研究;杨瑞馥、毕玉晶和高宏对实验进行了设计;常宇骁参与全部实验操作并分析了实验数据;李亨参与培养组学实验操作;丁磊收集了样本并对临床数据进行了分析;杨超参与了生物信息学分析;潘志远和韩妮对相关文献进行了调研;崔玉军和智发朝指导了实验;常宇骁、毕玉晶和杨瑞馥撰写并修改了文章。

## Compliance with ethics guidelines

Yuxiao Chang, Xiang Li, Lei Ding, Chao Yang, Zhiyuan Pan, Ni Han, Yujun Cui, Fachao Zhi, Ruifu Yang, Hong Gao, and Yujing Bi declare that they have no conflict of interest or financial conflicts to disclose.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.eng.2021.03.028>.

## References

- [1] Louis P, Hold GL, Flint HJ. The gut microbiota, bacterial metabolites and colorectal cancer. *Nat Rev Microbiol* 2014;12(10):661–72.
- [2] Peters BA, Dominianni C, Shapiro JA, Church TR, Wu J, Miller G, et al. The gut microbiota in conventional and serrated precursors of colorectal cancer. *Microbiome* 2016;4(1):69.
- [3] Sun J, Kato I. Gut microbiota, inflammation and colorectal cancer. *Genes Dis* 2016;3(2):130–43.
- [4] Yu J, Feng Q, Wong SH, Zhang D, Liang QY, Qin Y, et al. Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. *Gut* 2017;66(1):70–8.
- [5] Yang Y, Weng W, Peng J, Hong L, Yang L, Toiyama Y, et al. *Fusobacterium nucleatum* increases proliferation of colorectal cancer cells and tumor development in mice by activating Toll-like receptor 4 signaling to nuclear factor- $\kappa$ B, and up-regulating expression of microRNA-21. *Gastroenterology* 2017;152(4):851–66.e24.
- [6] Long X, Wong CC, Tong L, Chu ESH, Ho Szeto C, Go MYY, et al. *Peptostreptococcus anaerobius* promotes colorectal carcinogenesis and modulates tumour immunity. *Nat Microbiol* 2019;4(12):2319–30.
- [7] Barboza M, Sela DA, Pirim C, LoCascio RG, Freeman SL, German JB, et al. Glycoprofiling bifidobacterial consumption of galacto-oligosaccharides by mass spectrometry reveals strain-specific, preferential consumption of glycans. *Appl Environ Microbiol* 2009;75(23):7319–25.
- [8] Lopez-Siles M, Khan TM, Duncan SH, Harmsen HJM, Garcia-Gil LJ, Flint HJ. Cultured representatives of two major phylogroups of human colonic *Faecalibacterium prausnitzii* can utilize pectin, uronic acids, and host-derived substrates for growth. *Appl Environ Microbiol* 2012;78(2):420–8.
- [9] Wilson ID, Nicholson JK. Gut microbiome interactions with drug metabolism, efficacy, and toxicity. *Transl Res* 2017;179:204–22.
- [10] Lloyd-Price J, Mahurkar A, Rahnavard G, Crabtree J, Orvis J, Hall AB, et al. Strains, functions and dynamics in the expanded Human Microbiome Project. *Nature* 2017;550(7674):61–6.
- [11] Scholz M, Ward DV, Pasolli E, Tolio T, Zolfo M, Asnicar F, et al. Strain-level microbial epidemiology and population genomics from shotgun metagenomics. *Nat Methods* 2016;13(5):435–8.
- [12] Lagier JC, Hugon P, Khelaifia S, Fournier PE, La Scola B, Raoult D. The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. *Clin Microbiol Rev* 2015;28(1):237–64.
- [13] Sorbara MT, Littmann ER, Fontana E, Moody TU, Kohout CE, Gjonbalaj M, et al. Functional and genomic variation between human-derived isolates of Lachnospiraceae reveals inter- and intra-species diversity. *Cell Host Microbe* 2020;28(1):134–46.e4.
- [14] Allali I, Delgado S, Marron PI, Astudillo A, Yeh JJ, Ghazal H, et al. Gut microbiome compositional and functional differences between tumor and non-tumor adjacent tissues from cohorts from the US and Spain. *Gut Microbes* 2015;6(3):161–72.
- [15] Nakatsu G, Li X, Zhou H, Sheng J, Wong SH, Wu WKK, et al. Gut mucosal microbiome across stages of colorectal carcinogenesis. *Nat Commun* 2015;6(1):8727.
- [16] Ingala MR, Simmons NB, Wultsch C, Krampis K, Speer KA, Perkins SL. Comparing microbiome sampling methods in a wild mammal: fecal and intestinal samples record different signals of host ecology, evolution. *Front Microbiol* 2018;9:803.
- [17] Didelot X, Falush D. Inference of bacterial microevolution using multilocus sequence data. *Genetics* 2007;175(3):1251–66.
- [18] Jauregui F, Landraud L, Passet V, Diancourt L, Frapy E, Guigon G, et al. Phylogenetic and genomic diversity of human bacteremic *Escherichia coli* strains. *BMC Genomics* 2008;9(1):560.
- [19] Schuijt TJ, Lankelma JM, Scicluna BP, de Sousa e Melo F, Roelofs JJTH, de Boer JD, et al. The gut microbiota plays a protective role in the host defence against pneumococcal pneumonia. *Gut* 2016;65(4):575–83.
- [20] Tenaillon O, Skurnik D, Picard B, Denamur E. The population genetics of commensal *Escherichia coli*. *Nat Rev Microbiol* 2010;8(3):207–17.
- [21] Caugant DA, Levin BR, Selander RK. Genetic diversity and temporal variation in the *E. coli* population of a human host. *Genetics* 1981;98(3):467–90.
- [22] Bian Y, Du Y, Wang R, Chen N, Du X, Wang Y, et al. A comparative study of HAMSCs/HBMSCs transwell and mixed coculture systems. *IUBMB Life* 2019;71(7):1048–55.
- [23] Croxen MA, Law RJ, Scholz R, Keeney KM, Wlodarska M, Finlay BB. Recent advances in understanding enteric pathogenic *Escherichia coli*. *Clin Microbiol Rev* 2013;26(4):822–80.
- [24] Miceňková L, Bosák J, Vrba M, Ševčíková A, Šmajš D. Human extraintestinal pathogenic *Escherichia coli* strains differ in prevalence of virulence factors, phylogroups, and bacteriocin determinants. *BMC Microbiol* 2016;16(1):218.
- [25] Lindstedt BA, Finton MD, Porcellato D, Brandal LT. High frequency of hybrid *Escherichia coli* strains with combined Intestinal Pathogenic *Escherichia coli* (IPEC) and Extraintestinal Pathogenic *Escherichia coli* (ExPEC) virulence factors isolated from human faecal samples. *BMC Infect Dis* 2018;18(1):544.
- [26] Raisch J, Buc E, Bonnet M, Sauvanet P, Vazeille E, de Vallée A, et al. Colon cancer-associated B2 *Escherichia coli* colonize gut mucosa and promote cell proliferation. *World J Gastroenterol* 2014;20(21):6560–72.
- [27] Martínez-Medina M, García-Gil LJ. *Escherichia coli* in chronic inflammatory bowel diseases: an update on adherent invasive *Escherichia coli* pathogenicity. *World J Gastrointest Pathophysiol* 2014;5(3):213–27.
- [28] Camprubi-Font C, Ewers C, Lopez-Siles M, Martínez-Medina M. Genetic and phenotypic features to screen for putative adherent-invasive *Escherichia coli*. *Front Microbiol* 2019;10:108.
- [29] Prorok-Hamon M, Friswell MK, Alswied A, Roberts CL, Song F, Flanagan PK, et al. Colonic mucosa-associated diffusely adherent *afaC+* *Escherichia coli* expressing *lpfA* and *pks* are increased in inflammatory bowel disease and colon cancer. *Gut* 2014;63(5):761–70.
- [30] Suresh A, Ranjan A, Jadhav S, Hussain A, Shaik S, Alam M, et al. Molecular genetic and functional analysis of *pks*-harboring, extra-intestinal pathogenic *Escherichia coli* from India. *Front Microbiol* 2018;9:2631.
- [31] Nougayrède JP, Homburg S, Taieb F, Boury M, Brzuszkiewicz E, Gottschalk G, et al. *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. *Science* 2006;313(5788):848–51.
- [32] Gur C, Ibrahim Y, Isaacson B, Yamin R, Abed J, Gamliel M, et al. Binding of the Fap2 protein of *Fusobacterium nucleatum* to human inhibitory receptor TIGIT protects tumors from immune cell attack. *Immunity* 2015;42(2):344–55.
- [33] Kantele A, Lääveri T, Mero S, Häkkinen IMK, Kirveskari J, Johnston BD, et al. Despite predominance of uropathogenic/extraintestinal pathotypes among travel-acquired extended-spectrum  $\beta$ -lactamase-producing *Escherichia coli*, the most commonly associated clinical manifestation is travelers' diarrhea. *Clin Infect Dis* 2020;70(2):210–8.

- [34] Terlizzi ME, Gribaudo G, Maffei ME. UroPathogenic *Escherichia coli* (UPEC) infections: virulence factors, bladder responses, antibiotic, and non-antibiotic antimicrobial strategies. *Front Microbiol* 2017;8:1566.
- [35] Wijetunge DSS, Gongati S, DebRoy C, Kim KS, Couraud PO, Romero IA, et al. Characterizing the pathotype of neonatal meningitis causing *Escherichia coli* (NMEC). *BMC Microbiol* 2015;15(1):211.
- [36] Tyakht AV, Manolov AI, Kanygina AV, Ischenko DS, Kovarsky BA, Popenko AS, et al. Genetic diversity of *Escherichia coli* in gut microbiota of patients with Crohn's disease discovered using metagenomic and genomic analyses. *BMC Genomics* 2018;19(1):968.
- [37] Chang PH, Pan YP, Fan CW, Tseng WK, Huang JS, Wu TH, et al. Pretreatment serum interleukin-1 $\beta$ , interleukin-6, and tumor necrosis factor- $\alpha$  levels predict the progression of colorectal cancer. *Cancer Med* 2016;5(3):426–33.
- [38] Ray AL, Berggren KL, Restrepo Cruz S, Gan GN, Beswick EJ. Inhibition of MK2 suppresses IL-1 $\beta$ , IL-6, and TNF- $\alpha$ -dependent colorectal cancer growth. *Int J Cancer* 2018;142(8):1702–11.
- [39] Spurbeck RR, Dinh PC, Walk ST, Stapleton AE, Hooton TM, Nolan LK, et al. *Escherichia coli* isolates that carry *vat*, *fyuA*, *chuA*, and *yfcV* efficiently colonize the urinary tract. *Infect Immun* 2012;80(12):4115–22.
- [40] Rolhion N, Barnich N, Claret L, Darfeuille-Michaud A. Strong decrease in invasive ability and outer membrane vesicle release in Crohn's disease-associated adherent-invasive *Escherichia coli* strain LF82 with the *yfgL* gene deleted. *J Bacteriol* 2005;187(7):2286–96.
- [41] Alp D, Kuleaşan H, Korkut Altıntaş A. The importance of the S-layer on the adhesion and aggregation ability of lactic acid bacteria. *Mol Biol Rep* 2020; 47(5):3449–57.
- [42] Tapader R, Bose D, Basu P, Mondal M, Mondal A, Chatterjee NS, et al. Role in proinflammatory response of YghJ, a secreted metalloprotease from neonatal septicemic *Escherichia coli*. *Int J Med Microbiol* 2016;306 (7):554–65.
- [43] Tapader R, Bose D, Dutta P, Das S, Pal A, Raffatellu M. SslE (YghJ), a cell-associated and secreted lipoprotein of neonatal septicemic *Escherichia coli*, induces Toll-like receptor 2-dependent macrophage activation and proinflammation through NF- $\kappa$ B and MAP kinase signaling. *Infect Immun* 2018;86(9):e00399-18.
- [44] Ahern P, Faith J, Gordon J. Mining the human gut microbiota for effector strains that shape the immune system. *Immunity* 2014;40(6):815–23.
- [45] Bron PA, van Baarlen P, Kleerebezem M. Emerging molecular insights into the interaction between probiotics and the host intestinal mucosa. *Nat Rev Microbiol* 2012;10(1):66–78.