

## Atypical enteropathogenic *E. coli* are associated with disease activity in ulcerative colitis

Maximilian Baumgartner<sup>a</sup>, Rebecca Zirnbauer<sup>b</sup>, Sabine Schlager<sup>c</sup>, Daniel Mertens<sup>a</sup>, Nikolaus Gasche<sup>d</sup>, Barbara Sladek<sup>d</sup>, Craig Herbold<sup>e</sup>, Olga Bochkareva<sup>f</sup>, Vera Emelianenko<sup>f</sup>, Harald Vogelsang<sup>a</sup>, Michaela Lang<sup>a,g</sup>, Anton Klotz<sup>a</sup>, Birgit Moik<sup>c</sup>, Athanasios Makristathis<sup>e,h</sup>, David Berry<sup>e,g</sup>, Stefanie Dabsch<sup>a</sup>, Vineeta Khare<sup>a</sup>, and Christoph Gasche<sup>a</sup>

<sup>a</sup>Division of Gastroenterology and Hepatology, Department of Internal Medicine 3, Medical University of Vienna, Vienna, Austria; <sup>b</sup>Division of Visceral Surgery, Department of General Surgery, Medical University of Vienna, Vienna, Austria; <sup>c</sup>National Reference Laboratory for Escherichia coli, Austrian Agency for Health and Food Safety, Graz, Austria; <sup>d</sup>Biome Diagnostics GmbH, Vienna, Austria; <sup>e</sup>Joint Microbiome Facility of the Medical University of Vienna and the University of Vienna, Vienna, Austria; <sup>f</sup>Institute of Science and Technology Austria, Klosterneuburg, Austria; <sup>g</sup>Centre for Microbiology and Environmental Systems Science, Department of Microbiology and Ecosystem Science, Division of Microbial Ecology, University of Vienna, Vienna, Austria; <sup>h</sup>Division of Microbiology, Department of Laboratory Medicine, Medical University of Vienna, Vienna, Austria

### ABSTRACT

With increasing urbanization and industrialization, the prevalence of inflammatory bowel diseases (IBDs) has steadily been rising over the past two decades. IBD involves flares of gastrointestinal (GI) inflammation accompanied by microbiota perturbations. However, microbial mechanisms that trigger such flares remain elusive. Here, we analyzed the association of the emerging pathogen atypical enteropathogenic *E. coli* (aEPEC) with IBD disease activity. The presence of diarrheagenic *E. coli* was assessed in stool samples from 630 IBD patients and 234 age- and sex-matched controls without GI symptoms. Microbiota was analyzed with 16S ribosomal RNA gene amplicon sequencing, and 57 clinical aEPEC isolates were subjected to whole-genome sequencing and in vitro pathogenicity experiments including biofilm formation, epithelial barrier function and the ability to induce pro-inflammatory signaling. The presence of aEPEC correlated with laboratory, clinical and endoscopic disease activity in ulcerative colitis (UC), as well as microbiota dysbiosis. In vitro, aEPEC strains induce epithelial p21-activated kinases, disrupt the epithelial barrier and display potent biofilm formation. The effector proteins *espV* and *espG2* distinguish aEPEC cultured from UC and Crohn's disease patients, respectively. *EspV*-positive aEPEC harbor more virulence factors and have a higher pro-inflammatory potential, which is counteracted by 5-ASA. aEPEC may tip a fragile immune-microbiota homeostasis and thereby contribute to flares in UC. aEPEC isolates from UC patients display properties to disrupt the epithelial barrier and to induce pro-inflammatory signaling in vitro.

### ARTICLE HISTORY

Received 25 July 2022  
Revised 19 October 2022  
Accepted 28 October 2022



### KEYWORDS

Ulcerative colitis (UC); Crohn's disease (CD); inflammatory bowel disease (IBD); 5-ASA; p21-activated kinase (PAK); enteropathogenic *E. coli* (EPEC); Escherichia coli (*E. coli*); effector proteins; virulence factors; bacterial-epithelial interaction; microbiome


## Introduction

Ulcerative colitis (UC) and Crohn's disease (CD) are the most prevalent forms of inflammatory bowel diseases (IBD) and affect 0.5 – 1% of the Western population. IBD is characterized by chronically remittent inflammation of the gastrointestinal tract, associated with abdominal pain, diarrhea, intestinal mucosal ulceration and anemia. The molecular pathogenesis of IBD involves

a complex interplay of host genetics, aberrant mucosal immune response, epithelial barrier dysfunction, gut microbiome and environmental triggers that interfere with such. Genome-wide association studies have identified more than 240 risk loci which have been linked to an aberrant immune-microbiota homeostasis and intestinal barrier dysfunction.<sup>1,2</sup> IBD prevalence increases in an industrialized environment, more specific with

**CONTACT** Christoph Gasche  [christoph.gasche@meduniwien.ac.at](mailto:christoph.gasche@meduniwien.ac.at)  Division of Gastroenterology and Hepatology, Department of Internal Medicine III, Medical University of Vienna, Währinger Gürtel 18-20, Vienna, A-1090 Austria

Study concept and design: M.B., V.K. C.G., and M.B. drafted the manuscript with input from all authors. R.Z. performed multiplex qPCR analysis, D.M., R.Z., and M. B. established the clinical database, S.S. and B.M. cultivated the aEPEC strains, N.G. and B.S., provided the age- and sex-matched control samples without GI symptoms, M.B., C.H., O.B. V.E., and D.B. performed bioinformatics, H.V. provided IBD patient's stool samples, M.L., C.G., S.D., and M.B. wrote the ethics applications, and V.K., C.G., and S.D. supervised the project. V.K. and C.G. critically revised the manuscript. C.G. acquired funding.

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/19490976.2022.2143218>

© 2022 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

the use of food additives and pharmaceuticals that interfere with microbiota or barrier function. Emulsifiers, titanium dioxide, ethylenediaminetetraacetate acid (EDTA), NSAR and antibiotics serve as examples.<sup>3–6</sup>

The microbiota of IBD patients is characterized by reduced diversity and temporal instability.<sup>7</sup> During periods of disease flares, it shifts toward dysbiosis with a reduced abundance of short chain fatty acids producing obligate anaerobes such as *F. prausnitzii* and overgrowth of facultative anaerobes, particularly *E. coli*.<sup>8,9</sup> However, the exact sequelae and mechanisms connecting dysbiosis with flares in IBD disease activity remain elusive. IBD patients harbor *E. coli* isolates that primarily belong to B2 and D phylogroups with virulence factors that have originally been described in extraintestinal pathogenic *E. coli*.<sup>10–12</sup> Efficient horizontal gene transfer of pathogenicity islands (PAI) and high genetic plasticity in chromosomes and plasmids facilitate the rapid adaptation of *E. coli* to various ecological niches.<sup>13,14</sup> *E. coli* thrives in areas of ulcerations and its DNA has been detected in 80% of CD patient's granulomas.<sup>15,16</sup> Adherent-invasive *E. coli* in CD can replicate within macrophages and induce TNF- $\alpha$  secretion in vitro.<sup>17</sup> UC-associated *E. coli* were shown to potentiate intestinal inflammation in vivo.<sup>18,19</sup> Large-scale adoption of multiplex-PCR-based GI pathogen panels identified enteric pathogens, including attaching and effacing *E. coli* (AEEC), being connected to inflammatory flares in IBD.<sup>20,21</sup>

AEEC are diarrheagenic *E. coli* and possess the PAI locus of enterocyte effacement (LEE) which harbors a type three secretion system (T3SS), genes coding for the intimin protein (*eae*), its translocated receptor (*tir*), together with regulators, chaperones and other effector proteins such as *espG*<sup>22</sup> that are secreted into host cells. Their characteristic histopathological attaching and effacing lesions in intestinal epithelium result from cytoskeletal rearrangements due to binding of the adhesin intimin and translocated *tir*. Atypical enteropathogenic *E. coli* (aEPEC) are defined by possessing LEE, but lacking the EHEC (enterohemorrhagic *E. coli*) specific virulence factor Shiga toxin (*stx*) and the *E. coli* adherence factor bundle forming pili (*bfp*) of typical EPEC (tEPEC), which is necessary for localized adherence.<sup>23</sup> aEPEC are a heterologous group of organisms that developed by repeated acquisition of LEE variants into different chromosomal

backgrounds.<sup>24</sup> Depending on O (somatic) and H (flagellar) antigens, EPEC can be serotyped with 12 classic O-groups originally recognized by the World Health Organization. Over 80% of aEPEC reported in the literature belong to non-classical EPEC serogroups and more than one-quarter is O non-typable.<sup>25</sup> Clinically relevant EPEC serotypes have been highlighted in Supplementary Table 1. While EHEC and tEPEC are strongly associated with diarrhea, aEPEC is also prevalent in asymptomatic healthy individuals.<sup>26,27</sup> Lacking Shiga toxin and the ability for localized adherence, aEPEC pathogenicity seems to be dependent on effector protein repertoire and host susceptibility. T3SS effector proteins are secreted into host cells and alter a variety of pathways. Depending on the secretome composition, the net effect can be either pro- or anti-inflammatory.<sup>28</sup>

The majority of aEPEC T3SS effector proteins are not located on the LEE and their biological function remains poorly understood.<sup>29</sup> They cluster in PAI surrounded by transposase-like genes, suggesting horizontal acquisition.<sup>30</sup> The LEE-encoded effector protein *espG* has been shown to activate P-21 activated kinases (PAK).<sup>31</sup> PAKs are serine/threonine kinase effectors of small Rho GTPases Rac1/Cdc42 and orchestrate signaling cascades, involved in cytoskeletal reorganization, cell migration, wound healing, intestinal crypt homeostasis and innate immune response.<sup>32</sup> PAK 1/2 kinases are involved in host–pathogen interactions and are shown to be central to microbial infections. Bacterial pathogens and their virulent effector proteins hijack host cellular signaling pathways in which PAK1 is a key player.<sup>33,34</sup> Overactivation of PAK1 and PAK2 has been implicated as important drivers of colitis, providing a potential link between aEPEC infection and IBD flares.<sup>35–37</sup>

Here, we systematically studied the prevalence of diarrheagenic *E. coli* in IBD patients and age- and sex-matched controls without GI symptoms. We further investigated microbiota composition in aEPEC-positive UC patients and performed whole-genome sequencing of clinical aEPEC isolates, including analysis of non-LEE-effectors and strain phylogeny. Thereby, the present study enhances the understanding of this emerging opportunistic pathogen and provides potential opportunities for secondary prevention in UC.

## Material and methods

### Screening for diarrheagenic *E. coli*

A random subset of IBD patient's stool samples collected between 2012 and 2017 for calprotectin analysis (BÜHLMANN fCAL ELISA, Switzerland) at the Department of Internal Medicine III, Medical University of Vienna, Austria (n = 630 patients) were screened for the presence of aEPEC, tEPEC, EIEC, ETEC, EAEC or EHEC using a multiplex qPCR-based approach. Twenty-eight patients had longitudinal samples. Only the earliest time point was used for analysis of prevalence and clinical parameters to prevent bias. DNA from an age- and sex-matched control cohort (n = 234 subjects) without GI symptoms was provided by the commercial gut microbiome testing company myBioma (Austria). Calprotectin comparison between groups was done with the Mann–Whitney *U* test, prevalences were compared using Fisher's exact test. D'Agostino & Pearson omnibus normality test was applied prior to paired *t* test with longitudinal calprotectin values. The average time between longitudinal sample points was 9.25 months. All *p*-values from statistical tests in this study are two-tailed. For additional information see supplementary methods.

### Analysis of microbiota composition

16S rRNA gene amplicon sequencing of n = 25 aEPEC-pos/aEPEC-neg UC stool samples was performed as described previously. Briefly, the standard Illumina protocol and MiSeq technology were applied followed by amplicon sequence variant (ASV) analysis with DADA2<sup>38</sup> and modified Rhea scripts.<sup>39</sup> For taxonomic classification, SINA version 1.6.1 with the SILVA database SSU Ref NR 99 release 138 was used with default parameters. Differential abundance ASVs were analyzed using DESeq2.<sup>40</sup> Raw 16S rRNA amplicon sequencing data has been submitted to ncbi under the accession number PRJNA902016.

### Isolation of AEEC from IBD stool samples

Sixty-two stool samples that showed positivity in intimin (*eae*) PCR were sent to the Austrian Agency for Health and Food Safety (AGES, Austria) for isolation of AEEC, with a success rate of 43,5%. Bacterial isolation was based on colony picking

from selective agar followed by performing *eae* PCR first from pools, and if positive from single colonies. The procedure was stopped at 50 examined colonies per sample. Isolated AEEC were O-an H-serotyped by agglutination. Additionally, 10 aEPEC strains isolated from outbreaks of diarrheagenic disease were included in the analysis, and 20 strains isolated from healthy children were ordered from the Statens Serum Institut (Denmark). A list of the 57 strains used in this study can be found in Supplementary Table 2.

### In vivo AEEC pathogenicity experiments

Trans epithelial electrical resistance (TEER) experiments were performed using Caco-2 monolayers. Primary human colon epithelial cells (HCEC-1CT) were used for the assessment of pro-inflammatory signaling (IL-8 secretion and p21-kinase expression) induced by aEPEC. Pairwise comparison was performed with the Mann–Whitney *U* test and ANOVA with Dunn's multiple comparison test for comparing multiple groups. Additional information on experimental setup and cell culture media/conditions can be found in the supplementary method section.

### In vitro biofilm formation assay

Fifty-seven *E. coli* isolates were grown on MacConkey agar for 24 hours under aerobic or anaerobic conditions, using Anaerobox and AnaeroGen sachets (Thermo Scientific, Oxoid). Single colonies were inoculated in 5 ml brain heart infusion (BHI, 37 g/L) medium with supplements (5 g/L yeast extract, 1 g/L NaHCO<sub>3</sub>, 1 g/L L-cysteine, 1 mg/L vitamin K1, 5 mg/L hemin) or in LB medium and grown under aerobic or anaerobic conditions for 6 hours at 37°C. Bacterial cells were diluted to an OD<sub>600</sub> = 0.05. 100 µl of cell suspension was transferred to the U-bottom polystyrene 96-well plates (Costar) in four technical replicates. Plates were incubated at 37°C for 48 hours under aerobic or anaerobic conditions. Supernatants were removed, bacterial biofilms were fixed with 150 µL BOUIN solution (0.9% picric acid, 9% formaldehyde and 5% acetic acid) for 15 min and washed three times with 190 µL PBS. For staining, 150 µL 0.1% crystal violet solution was added for 10 min and washed

three times with H<sub>2</sub>O. For biofilm quantification, crystal violet in dried plates was dissolved in 190 µL 30% acetic acid and the plate was placed on a shaker for 1 h. Absorbance of 1:5 dilutions was measured on an Anthos 2010 microplate reader at 595 nm and 405 nm reference wavelength. Pairwise comparison was performed with the Mann–Whitney *U* test and ANOVA with Dunn's multiple comparison test for comparing multiple groups.

### **Whole-genome sequencing and bioinformatic analysis**

Bacterial DNA was extracted using a phenol chloroform-based method. Whole-genome sequencing was performed using HiSeqV4 PE125 methodology. For genome assembly, the spades pipeline was used. Assemblies were submitted to NCBI for annotation. The CFSAN SNP pipeline was used with the *E. coli* reference genome O103:H2 12009 to construct an SNP matrix with the 57 strains from this study and 348 publicly available AEEC genomes of diverse pathotypes and one *E. albertii* genome. For phylogenomic maximum likelihood inference, IQ-TREE was applied with the best-fit model automatically selected by ModelFinder.<sup>41</sup> For pangenome analysis, the Roary pipeline was used with standard parameters, followed by Scoary for the identification of associations between all genes in the accessory genome and EspG2 and EspV positivity.<sup>42</sup> Pangenome composition was visualized with Phandango. To investigate the presence of known virulence factors, the VFDB database was used. For the detection of novel hypothetical secreted proteins and additional secretion systems, the EffectiveDB was applied. Pairwise comparison between EspG2-pos and EspV-pos genomes was performed with the Mann–Whitney *U* test, prevalences were compared using Fisher's exact test, with Bonferroni correction for multiple comparisons. Sequencing data and assemblies are publicly accessible at NCBI under the project number PRJNA528578. Additional information on bioinformatic analysis can be found in the supplementary method section.

### **Ethics statement**

The study was reviewed and approved by the ethics committee of the Medical University of Vienna (EK-Nr: 1522/2015). The study was conducted in

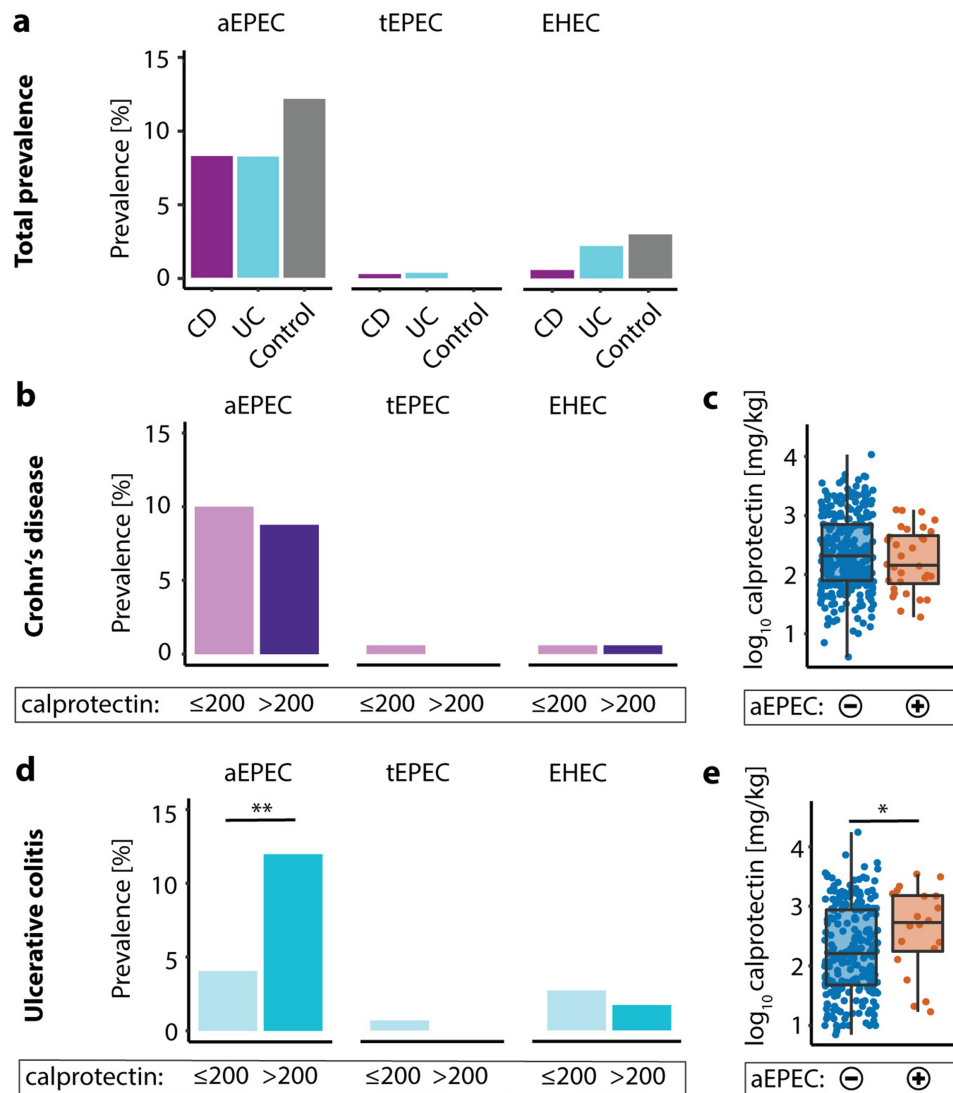
accordance with the ethical principles expressed in the Declaration of Helsinki and the requirements of applicable federal regulations.

## **Results**

### **Atypical EPEC correlate with disease activity in UC**

We first screened fecal samples from patients with UC (n = 274), CD (n = 356) and age- and sex-matched controls without GI symptoms (n = 234) for the presence of AEEC using a multiplex qPCR-based approach. EHEC and tEPEC were rare, with less than 3% and 0.4% prevalence in all cohorts, respectively. The diarrheagenic *E. coli* subtype enteroaggregative *E. coli* (EAEC) had less than 4% prevalence and enterotoxigenic *E. coli* (ETEC) as well as enteroinvasive *E. coli* (EIEC) were not detectable in our cohort. aEPEC, however, could be detected in approximately 10% of samples (Figure 1a). To determine the connection between active GI inflammation and presence of aEPEC, fecal calprotectin was analyzed in the same samples.<sup>43</sup> There was no association between calprotectin and aEPEC positivity in CD (Figure 1b and c). However, UC patients with GI inflammation had increased aEPEC prevalence compared to UC patients without GI inflammation (12% vs. 4%, *p* = .01, Figure 1d). aEPEC-positive (aEPEC-pos) UC patients exhibited median calprotectin values that were more than three times as high as aEPEC negative (aEPEC-neg) (625 vs. 162 mg/kg, *p* = .01, Figure 1e). Endoscopic and clinical disease activity were also higher in EPEC-pos UC patients, suggesting a link between aEPEC and flares of disease activity in UC (Table 1). In CD, there was no difference in endoscopic or clinical disease activity between aEPEC-pos and aEPEC-neg patients (Supplementary Table 2). Age, sex, disease extent, age of diagnosis, medication and smoking status were not associated with aEPEC in CD and UC (Table 1, Supplementary Table 2). Enterotoxigenic *E. coli* (EAEC) did not correlate with GI inflammation (Supplementary Figure 1).

Longitudinal analysis confirmed the link of aEPEC and GI inflammation in UC, as patients had lower calprotectin at aEPEC-neg points of time, which was not seen for CD patients (Figure 2a and b). Tracing aEPEC status over time showed transient phases of aEPEC positivity in IBD patients, suggesting re-infection or abundances below the limit of



**Figure 1.** aEPEC is more prevalent in UC patients with active disease. (a) Prevalence of aEPEC, tEPEC and EHEC in CD- (purple), UC-patients (blue) and age and sex matched controls without GI-symptoms (gray). (b) Prevalence of aEPEC, tEPEC and EHEC in CD with fecal calprotectin below (light purple) and above 200 mg/kg (dark purple). (c) Fecal calprotectin in aEPEC-neg (blue) and aEPEC-pos (Orange) CD patients. (d) Prevalence of aEPEC, tEPEC and EHEC in UC with fecal calprotectin below (light blue) and above 200 mg/kg (blue). (e) Fecal calprotectin in aEPEC-neg (blue) and aEPEC-pos (Orange) UC patients. Statistical analysis: (a,b,c) Fisher's exact text, (c,e)  $\log_{10}$  transformed y-axis, Mann-Whitney  $U$  test,  $n = 356$  CD, 274 UC and 234 controls; \* $p \leq .05$ ; \*\* $p \leq .01$ .

detection of a resident aEPEC (Figure 2c). 16S rRNA gene amplicon sequencing revealed reduced bacterial diversity as defined by Shannon index in aEPEC-pos UC patients compared to aEPEC-neg (Figure 2d). aEPEC-pos UC patients had increased abundances of amplicon sequencing variants (ASV) belonging to *Dialister*, *Haemophilus* and *Veillonella*. RDA-analysis showed a significant effect of aEPEC-positivity on microbiome composition, with fecal calprotectin following a similar gradient (Supplementary Figure 2a). ASV belonging to protein metabolizing *Acidaminococcus* were reduced in

aEPEC-pos UC. Furthermore, several ASV belonging to *Bacteroides* were reduced and one was increased (Figure 2e). Adjusting the DESeq2 model for GI inflammation confirmed an increase in ASV belonging to *Haemophilus* in aEPEC-pos UC and revealed an enrichment of ASV belonging to sulfate reducing *Bilophila* and several beneficial bacteria such as *Eubacterium* and *Subdoligranulum* (Supplementary Figure 2b and c). Overall, these findings support the concept that aEPEC are correlated with flares in disease activity and microbiota dysbiosis in UC.

**Table 1.** Clinical parameters of aEPEC-pos UC patients.

Cohort	Variables	DEC <sup>-</sup> patients (n = 550)	aEPEC <sup>+</sup> patients (n = 53)	p-value <sup>†</sup>
Ulcerative colitis	Total number	238	23	
	Sex (% female), n (%)	111 (47%)	8 (35%)	0.38
	Age at sampling [years]	43 [32–56]	41 [27–57]	0.52
	Age at diagnosis [years]	27 [23–41]	26 [19–33]	0.27
	Proctitis (E1), n (%)	23/168 (14%)	3/16 (19%)	0.34
	Left-sided colitis (E2), n (%)	72/168 (43%)	4/16 (25%)	
	Pancolitis (E3), n (%)	73/168 (43%)	9/16 (56%)	
	IBD related surgery, n (%)	21/238 (9%)	3/23 (13%)	0.45
	<b>Endoscopically inactive disease<sup>†</sup> (Mayo 0), n (%)</b>	<b>33/78 (42%)</b>	<b>2/8 (25%)</b>	<b>0.04*</b>
	<b>Mild endoscopic disease<sup>†</sup> (Mayo 1), n (%)</b>	<b>19/78 (24%)</b>	<b>1/8 (12%)</b>	
	<b>Moderate endoscopic disease<sup>†</sup> (Mayo 2), n (%)</b>	<b>13/78 (17%)</b>	<b>0/8 (0%)</b>	
	<b>Severe endoscopic disease<sup>†</sup> (Mayo 3), n (%)</b>	<b>13/78 (17%)</b>	<b>5/8 (63%)</b>	
	<b>Clinical disease activity[partial mayo score]</b>	<b>1 [0–3]</b>	<b>2 [1–5]</b>	<b>0.03*</b>
	<b>Fecal calprotectin [mg/kg]</b>	<b>162 [47–864]</b>	<b>625 [178–1677]</b>	<b>0.012*</b>
	C-reactive protein [mg/L]	0.25 [0.08–0.71]	0.32 [0.12–1.86]	0.23
	Non-antibody-based immunotherapy, n (%)	55/207 (27%)	6/21 (29%)	0.8
	Antibody-based immunotherapy, n (%)	51/207 (25%)	5/21 (24%)	1
	5-ASA, n (%)	176/207 (85%)	16/21 (76%)	0.34
	Corticosteroids, n (%)	21/207 (10%)	4/21 (19%)	0.26
	Probiotics, n (%)	22/207 (11%)	3/21 (14%)	0.71
Nicotine, n (%)	18/207 (9%)	1/21 (5%)	1	

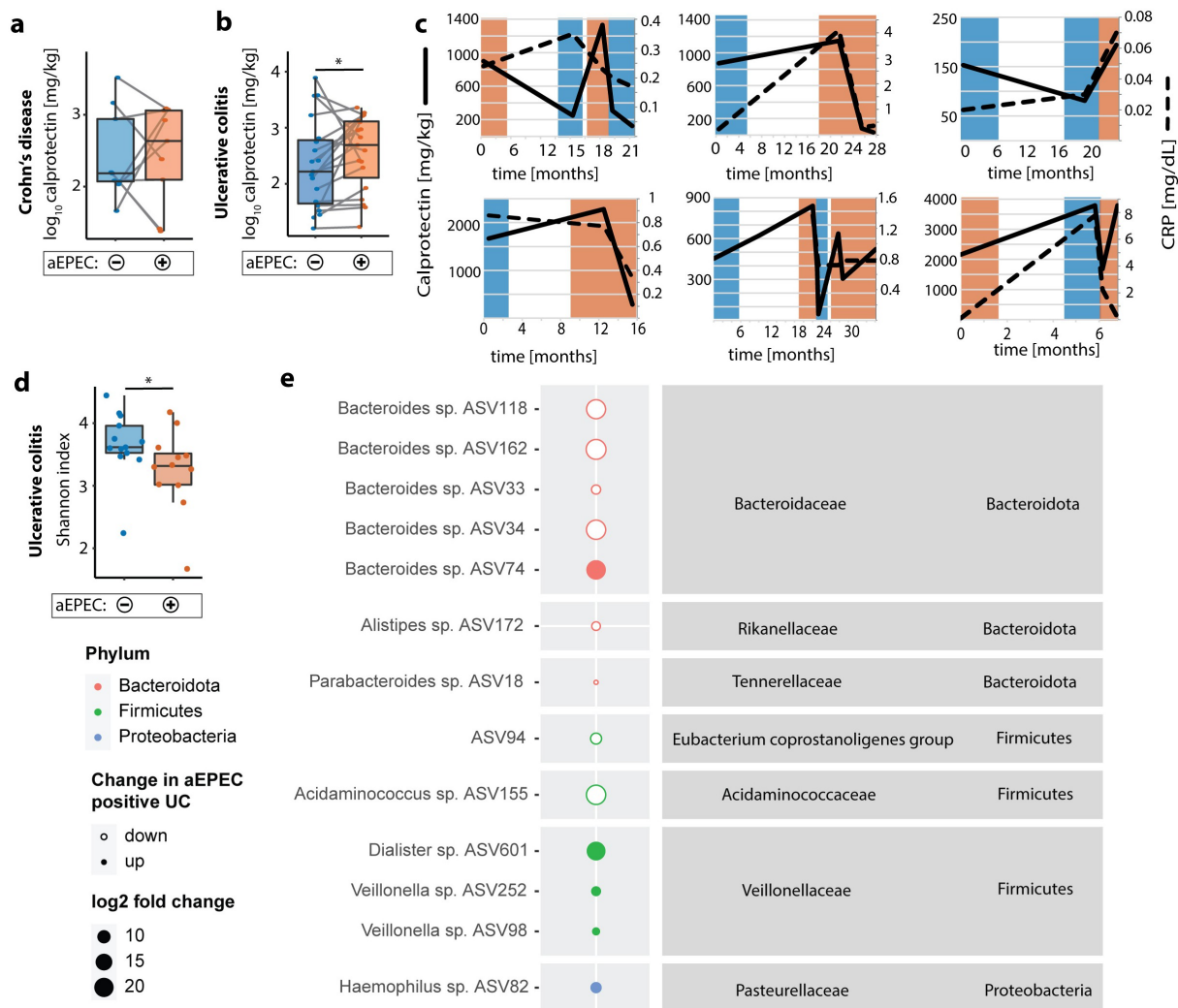
Notes: Values are presented as median with range in brackets for continuous variables or number and percentage in brackets for categorical variables. Percentages are calculated based on the actual number of patients in each group where the respective data was available. If data was not available for all subjects, the number of subjects for which the respective data was available is indicated after the backslash. The Mann–Whitney *U* test and two-sided Fisher's exact test were used to determine p-values for continuous and categorical variables, respectively. †aEPEC positive patients vs. DEC negative patients, \**p* ≤ 0.05.

Abbreviations: DEC, Diarrheagenic *E. coli*; aEPEC, atypical enteropathogenic *E. coli*.

### UC-associated aEPEC elicit a pro-inflammatory response in vitro

To investigate how aEPEC could trigger GI inflammation and if aEPEC from UC patients behave differently than aEPEC from CD patients, we performed in vitro pathogenicity experiments using strains isolated from IBD patient's fecal samples (n = 13 UC, n = 12 CD). Disruption of epithelial tight junction (TJ) barrier is a well-defined pathomechanism of tEPEC infection, thereby inducing diarrhea. To test for TJ barrier disruption, bacteria were co-cultivated with Caco-2 monolayers, and barrier function was assessed continuously using electric cell-substrate impedance sensing (ECIS) technology. The tEPEC-E2348/69 reference strain led to a steady decrease in transepithelial electrical resistance (TEER) with a ~50% drop after 3 h. aEPEC strains from CD and UC patients showed comparable TEER responses with an initial rise of barrier function which peaked at 2 h, followed by an abrupt drop as the infection progressed (Figure 3a). Recently, bacterial biofilm formation has been implicated in IBD pathophysiology.<sup>44</sup> aEPEC built strong biofilms in rich BHI media and under aerobic conditions. Biofilm formation was less pronounced with LB media without additional glucose and in anaerobic conditions. There was no difference between CD- and UC-associated aEPEC regarding biofilm

formation (Figure 3b). Depending on T3SS effector proteins, EPEC can induce either a pro- or anti-inflammatory epithelial response.<sup>24</sup> Thus, aEPEC strains were co-cultivated with immortalized human primary colon epithelial cells (HCEC-1CT), and IL-8 secretion was measured as a marker of pro-inflammatory signaling. In agreement with our clinical findings, UC-associated aEPEC elicited a stronger IL-8 response compared to CD-associated aEPEC (Figure 3c). tEPEC-E2348/69 attenuated IL-8 secretion, while *E. coli* K-12 induced IL-8 secretion at a comparable rate to UC-associated aEPEC (Supplementary Figure 3b). When comparing in vitro pathogenicity findings with aEPEC strains isolated from infectious diarrhea patients (n = 12) and healthy controls (n = 20), aEPEC from healthy controls led to a higher initial increase and less pronounced drop in TEER, as well as showing stronger biofilm formation under anaerobic conditions (Supplementary Figure 3a and c). *E. coli* K-12 had stronger biofilm formation than IBD-associated aEPEC under anaerobic conditions (Figure 3b). Compared to UC-associated aEPEC, aEPEC from infectious diarrhea patients had weaker biofilm formation in LB media without glucose (Supplementary Figure 3c). The majority of aEPEC strains isolated from IBD patients belonged to previously unrecognized non-classical aEPEC serotypes and there was no



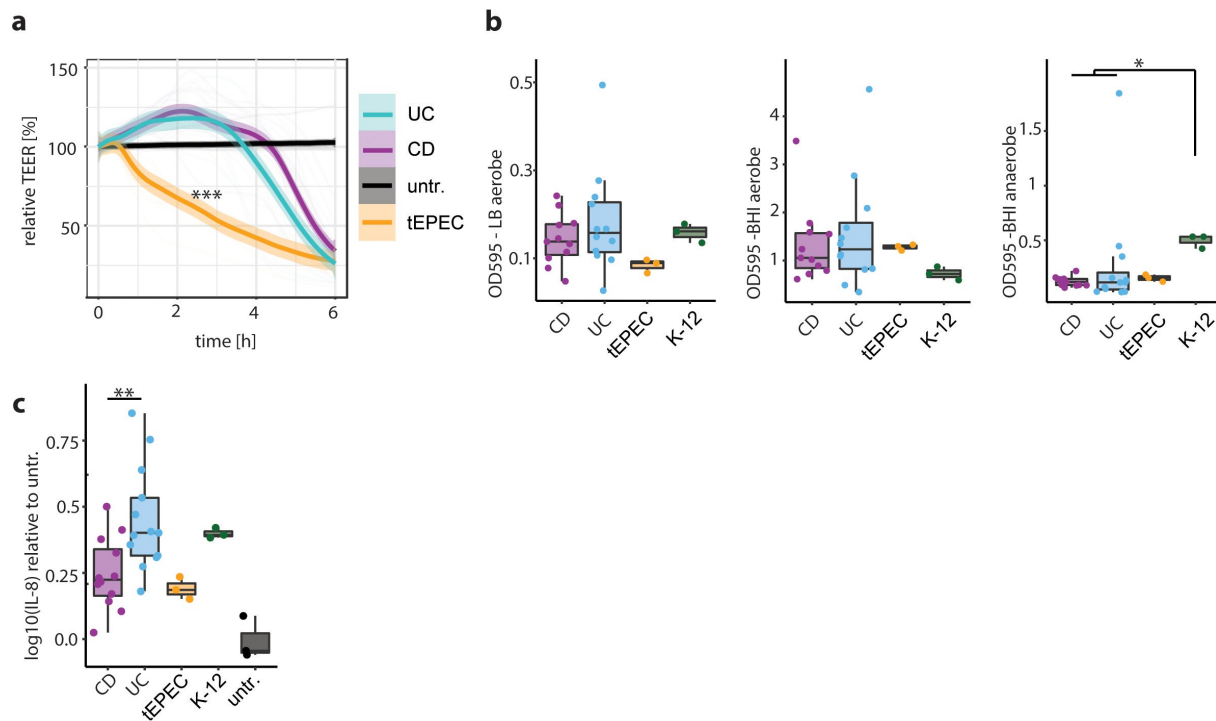
**Figure 2.** Longitudinal clinical parameters and microbiome analysis of aEPEC-pos UC patients. (a) Longitudinal fecal calprotectin in aEPEC-neg (blue) and aEPEC-pos (Orange) UC patients, samples from the same patient connected with gray lines. (b) Longitudinal fecal calprotectin in aEPEC-neg (blue) and aEPEC-pos (Orange) CD patients, samples from the same patient connected with gray lines. (c) Longitudinal trajectories of fecal calprotectin and serum c-reactive protein, with aEPEC-pos timepoints (Orange background) and aEPEC-neg timepoints (blue background). (d) UC patient's stool bacterial diversity represented by Shannon index, aEPEC-neg (blue) and aEPEC-pos (Orange). (e) DESeq2 analysis at ASV level, aEPEC-pos vs. aEPEC-neg UC patients, Size represents fold-change, full dots represent up-regulation, empty dots down-regulation. Significant findings ( $p < .05$  corrected for multiple comparisons) are shown. Statistical analysis: (a,b) Log<sub>10</sub> transformed y-axis, ratio paired t test,  $n = 9$  CD, 19 UC paired samples (d) Mann-Whitney  $U$  test, (d,e)  $n = 12$  aEPEC-pos and 13 aEPEC-neg UC patients; \* $p \leq .05$ .

association between serotype and disease cohort (Supplementary Tables 1, 3). Altogether, these data indicate that UC-associated aEPEC show virulent in vitro phenotypes, resembling biofilm formation, barrier dysfunction and inflammation.

### Non-LEE effector proteins *EspG2* and *EspV* distinguish aEPEC from CD and UC

To analyze the population structure, a maximum likelihood (ML) phylogeny was constructed using

a reference-based single nucleotide polymorphism matrix of aEPEC isolates from IBD patients, infectious diarrhea patients and healthy controls (total  $n = 57$ ), together with 348 publicly available AEEC genomes of diverse pathotypes and one *E. albertii* genome. The isolated strains were initially identified to be aEPEC based on PCR detection of *eae* but not *bfpA* or *stx*. Analysis of sequencing data revealed one supposed aEPEC isolate from the healthy cohort to possess *stx1*. The strain was reclassified as EHEC together with another strain



**Figure 3.** aEPEC strains from UC elicit a more pro-inflammatory in vitro response than CD. (a) Trans epithelial resistance of Caco-2 monolayers infected with aEPEC strains isolated from UC patients (blue), CD patients (purple) and reference strain tEPEC-E2348/69 (Orange), with untreated cells (black). (b) Biofilm formation assay of aEPEC strains isolated from UC patients (blue) and CD patients (purple), tEPEC-E2348/69 (Orange), *E. coli* K-12 (green) aerobic conditions with LB medium (left), aerobic conditions with BHI medium (middle) and anaerobic conditions with BHI medium (right). (c) IL-8 secretion of human primary colon epithelial cells (HCEC-1CT) infected with aEPEC strains isolated from UC patients (blue) and CD patients (purple), tEPEC-E2348/69 (Orange), *E. coli* K-12 (green) normalized to untreated cells (black). Statistical analysis: (a,b)  $n = 11$  CD, 12 UC, (c)  $n = 12$  CD, 13 UC, (a) two-way ANOVA with Tukey's multiple comparisons test, (b,c) Mann–Whitney  $U$  test; \* $p \leq .05$ ; \*\* $p \leq .01$ ; \*\*\* $p \leq .001$ .

from the healthy cohort which was reclassified as *E. albertii* after phylogenetic analysis (Supplementary Figure 4a). The AEEC genomes clustered in 14 clonal groups (CG) containing >5 isolates which were named based on their dominant Achtman multi-locus sequence type (Supplementary Figure 4a). aEPEC evolved through multiple LEE acquisition events via horizontal gene transfer.<sup>24</sup> To compare evolutionary history of LEE with the whole-genome evolution in the 57 isolated strains, ML trees were generated using aligned LEE encoded genes and a concatenated alignment of 2719 single-copy common genes (Supplementary Figure 4c). Comparison of the resulting trees points toward possible recombination events within the LEE (Extended Data Supplementary Figure 1). The LEE sequences clustered in three lineages, with the majority of strains belonging to the LEE1 and LEE3 lineages (Supplementary Figure 4d). LEE1 had more known non-LEE effector proteins than

LEE3 ( $p < .006$ , two-sided Mann–Whitney  $U$  test with Bonferroni correction). Ninety-one percent of CD-associated aEPEC belonged to LEE3 compared to 54% of UC-associated aEPEC (Supplementary Table 4, Fisher's exact test,  $p < .05$ ). There was no association between CG and disease cohort. Intimin comprises three major ( $\alpha$ ,  $\beta$  and  $\gamma$ ) and multiple minor subtypes (epsilon:  $\epsilon$ , iota:  $\iota$  and zeta:  $\zeta$ ) which have been linked to tissue tropism and evolutionary branches of EPEC.<sup>45,46</sup> The intimin subtypes of our isolated strains were scattered across disease cohorts (Supplementary Table 3). As effector protein composition could explain the different in vitro behavior of UC- and CD-associated aEPEC, an exploratory analysis of known non-LEE effector proteins was performed using recursive partitioning. EspV was more abundant among UC-associated aEPEC strains (61% vs. 8%, Fisher's exact:  $p < .05$ ), and EspG2 was more prevalent in aEPEC strains from CD patients (50% vs. 8%, Fisher's exact:  $p < .05$ ) (Figure 4a, Supplementary



Figure 5). Taken together, these findings suggest that distinct subgroups of aEPEC are detectable in UC vs. CD patients.

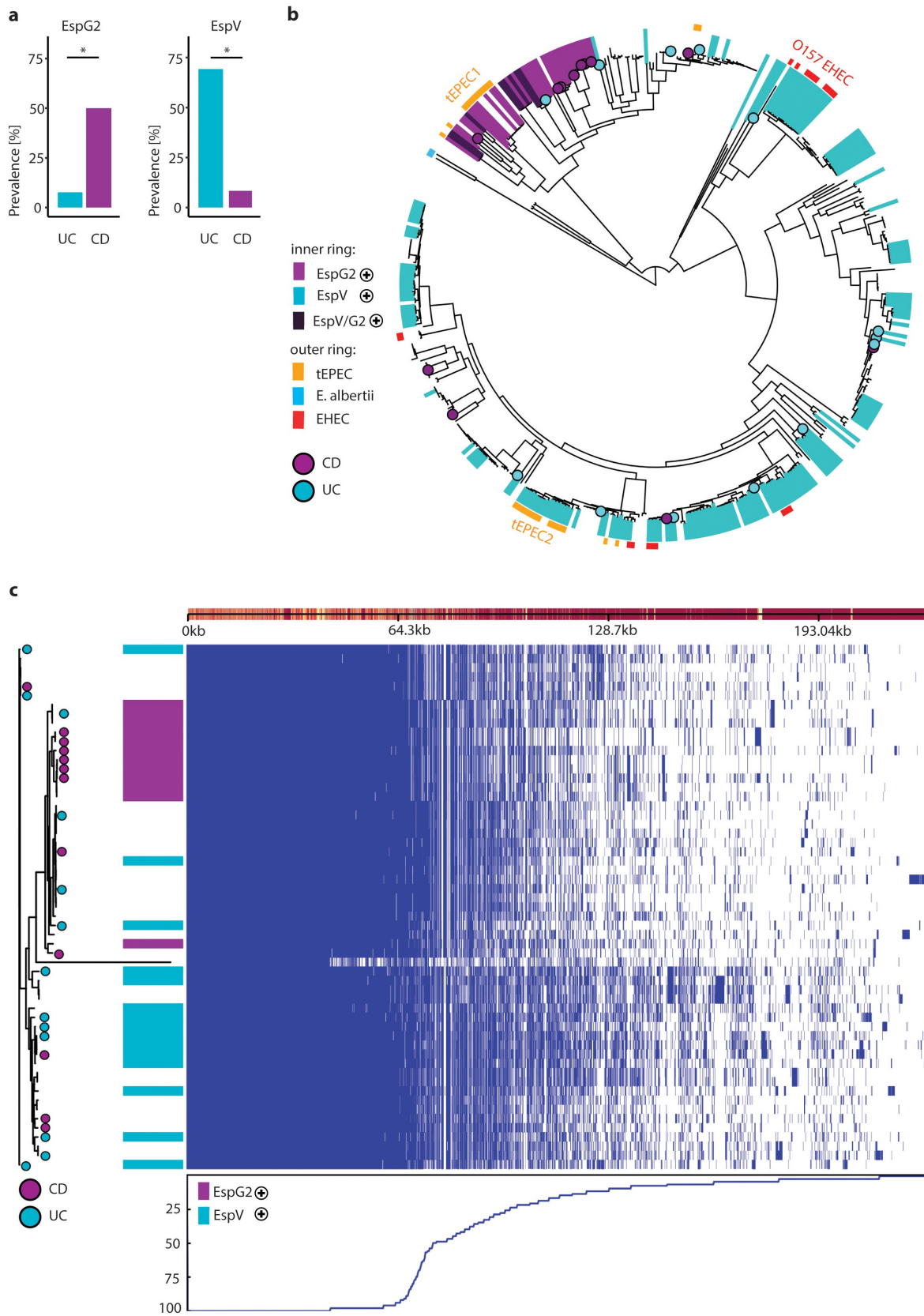
### ***EspV-positive aEPEC are more virulent than EspG2-positive aEPEC***

Typical EPEC evolved multiple times within *E. coli* through independent acquisition events of LEE PAI and the EAF plasmid. Strains have traditionally been classified into two major clades EPEC1 and EPEC2, which belong to phylogroups B2 and B1, respectively.<sup>47</sup> In humans, intimin  $\alpha$  is specifically expressed by EPEC1, and intimin  $\beta$  has primarily been associated with EPEC2.<sup>45</sup> EspG2 is a non-LEE encoded homolog of the LEE gene EspG and has been described in EPEC1.<sup>48</sup> Indeed, aEPEC that harbor EspG2 clustered together in close proximity to the ‘typical’ EPEC1 clade, which also includes tEPEC-E2348/69 (bootstrap: 100%), while EspV harboring aEPEC were scattered across the tree. The only clonal group among EspG2-pos aEPEC was CG526, while EspV-pos aEPEC strains were placed in six different clonal groups (Figure 4b, Supplementary Figure 4a). All EspG2-positive (EspG2-pos) aEPEC from the 57 isolated strains had the LEE3 lineage and all intimin  $\alpha$  possessing strains fell into this clade as well (Supplementary Figure 4d). A subgroup of EspG2-pos aEPEC were also EspV-positive (EspV-pos), however none of the 57 isolated strains possessed both genes. Eighty percent of EHEC genomes were EspV-pos compared to 40% of aEPEC strains, indicating a correlation with in vivo pathogenicity (Fisher’s exact:  $p < .001$ ). To investigate further genetic differences between EspG2-pos and EspV-pos aEPEC strains, we performed pangenome analysis of the 57 isolated strains using Roary. The number of genes in the resulting pangenome continues to increase non-asymptotically with each additionally added isolate, classifying it as an open pangenome (Extended Data Supplementary Figure 2). Twenty percent (3190) of the identified genes were ‘core genes’ present in more than 95% of the isolates. ‘Accessory genes’ represented 80% of the pangenome (12895 genes). A large proportion of the accessory genomes were present in fewer than 15% of the strains (10191/12895 genes), highlighting genomic diversity and plasticity of AEEC. Visualizing the

pangenome with phandango revealed fingerprint patterns of the accessory genome in EspG2-pos and EspV-pos clusters, hinting at an association of distinct genetic compositions with EspG2 and EspV positivity (Figure 4c). To identify genes associated with EspG2 and EspV we utilized the microbial pan-GWAS pipeline Scoary, which discovered 968 genes distinguishing EspG2-pos and EspV-pos strains ( $p < .05$  after Benjamini–Hochberg correction). Gene ontology analysis linked the genes with benzene-containing compound metabolic processes and cellular response to xenobiotic stimulus ( $p < 7.4E-09$ ). Among the genes co-occurring with EspV were several virulence factors, laminin-binding fimbriae (*elfD/G*), type 1 fimbriae (*fimH*) and multiple genes annotated as iron ABC transporter permeases (Extended data Table 1). The virulence factor database (VFDB) is a manually curated database for virulence factors of medically important pathogens. Using the VFDB, genomes of EspV-pos aEPEC were shown to possess more known virulence factors linked to adherence and invasion, as well as auto-transporters and toxins (Table 2). Seventy-five percent of EspV-pos aEPEC had the gene encoding for hemolysin E and none for EspG2-pos. Furthermore, the EffectiveDB pipeline was applied to discover novel T3SS effector proteins based on their N-terminal signal peptide and classify if strains possess a functioning type 4 and 6 secretion system (T4SS, T6SS). EspV-pos strains had more total predicted secreted proteins, T3SS effector proteins (median T3SS: 473 vs. 383,  $p < .003$ ) and T3SS chaperones. Sixty-seven percent of EspV-pos aEPEC had a functioning T6SS predicted with high confidence, compared to none of the EspG2-pos strains (Table 2). Overall, the genomic data indicate increased potential for virulence in EspV-pos aEPEC.

### ***aEPEC induce PAKs and 5-ASA counteracts their pro-inflammatory stimulus***

During infection, typical EPEC inactivates the innate immune response via various translocated effector proteins that prevent IKK-mediated phosphorylation of I $\kappa$ B and NF- $\kappa$ B, prior to TJ disruption.<sup>49</sup> This results in the phenotype of watery diarrhea with a rather weak inflammatory response. In our in vitro model using co-cultivation



**Figure 4.** Phylogeny of AEEC and characterization of EspG2 and EspV-carrying aEPEC. (a) Prevalence of non-LEE effectors EspG2 (left) and EspV (right) in aEPEC isolates from UC- (blue) and CD-patients (purple). (b) Midpoint rooted tree constructed with 405 AEEC genomes. Strains isolated from UC- (blue circle) and CD-patients (purple circle) labeled at the tip. Inner ring depicts the presence of EspG2 (purple), EspV (blue) or both (dark purple) in the genomes. Outer ring depicts pathotype; tEPEC (Orange), *E. albertii* (light blue)

**Table 2.** Genomic characteristics of clinical aEPEC isolates.

Database	Variables	EspV <sup>+</sup> aEPEC (n = 15)	EspG2 <sup>+</sup> aEPEC (n = 12)	p-value <sup>†</sup>
VFDB	<b>Adherence</b>	<b>28 (24–30)</b>	<b>21.5 (21–28)</b>	<b>0.001</b>
	<b>Autotransporter</b>	<b>4 (2–6)</b>	<b>3 (2–3)</b>	<b>0.01</b>
	<b>Invasion</b>	<b>2 (1–3)</b>	<b>3 (2–3)</b>	<b>0.034</b>
	Iron uptake	10 (0–18)	7 (7–21)	0.881
	<b>Toxin</b>	<b>1 (0–7)</b>	<b>0 (0–3)</b>	<b>0.003</b>
	Immune evasion	1 (0–2)	1 (0–2)	0.529
	Anti-phagocytosis	0 (0–1)	0 (0–1)	0.624
EffectiveDB	<b>Predicted secreted proteins</b>	<b>996 (808–1121)</b>	<b>827.5 (804–889)</b>	<b>0.0004</b>
	<b>Predicted T3SS effectors</b>	<b>473 (381–543)</b>	<b>383 (375–418)</b>	<b>0.003</b>
	<b>Conserved binding domains of T3SS chaperones</b>	<b>106 (94–117)</b>	<b>101.5 (95–106)</b>	<b>0.036</b>
	<b>Predicted T4SS effectors</b>	<b>118 (96–130)</b>	<b>102.5 (93–124)</b>	<b>0.015</b>
	Unique eukaryotic-like domains	28 (22–35)	29.5 (20–32)	0.435
	Nr. of proteins with eukaryotic-like domains	104 (65–131)	94.5 (59–104)	0.180
	High likelihood of functioning T3SS	15/15 (100%)	12/12 (100%)	1
	High likelihood of functioning T4SS	3/15 (20%)	2/12 (17%)	1
	<b>High likelihood of functioning T6SS</b>	<b>10/15 (67%)</b>	<b>0/12 (0%)</b>	<b>0.0004</b>

Notes: Values are presented as median with first and third quartiles in brackets for continuous variables or number and percentage in brackets for categorical variables. The Mann–Whitney *U* test and two-sided Fisher's exact test were used to determine p-values for continuous and categorical variables, respectively. <sup>†</sup>EspV positive vs. EspG2 positive strains.

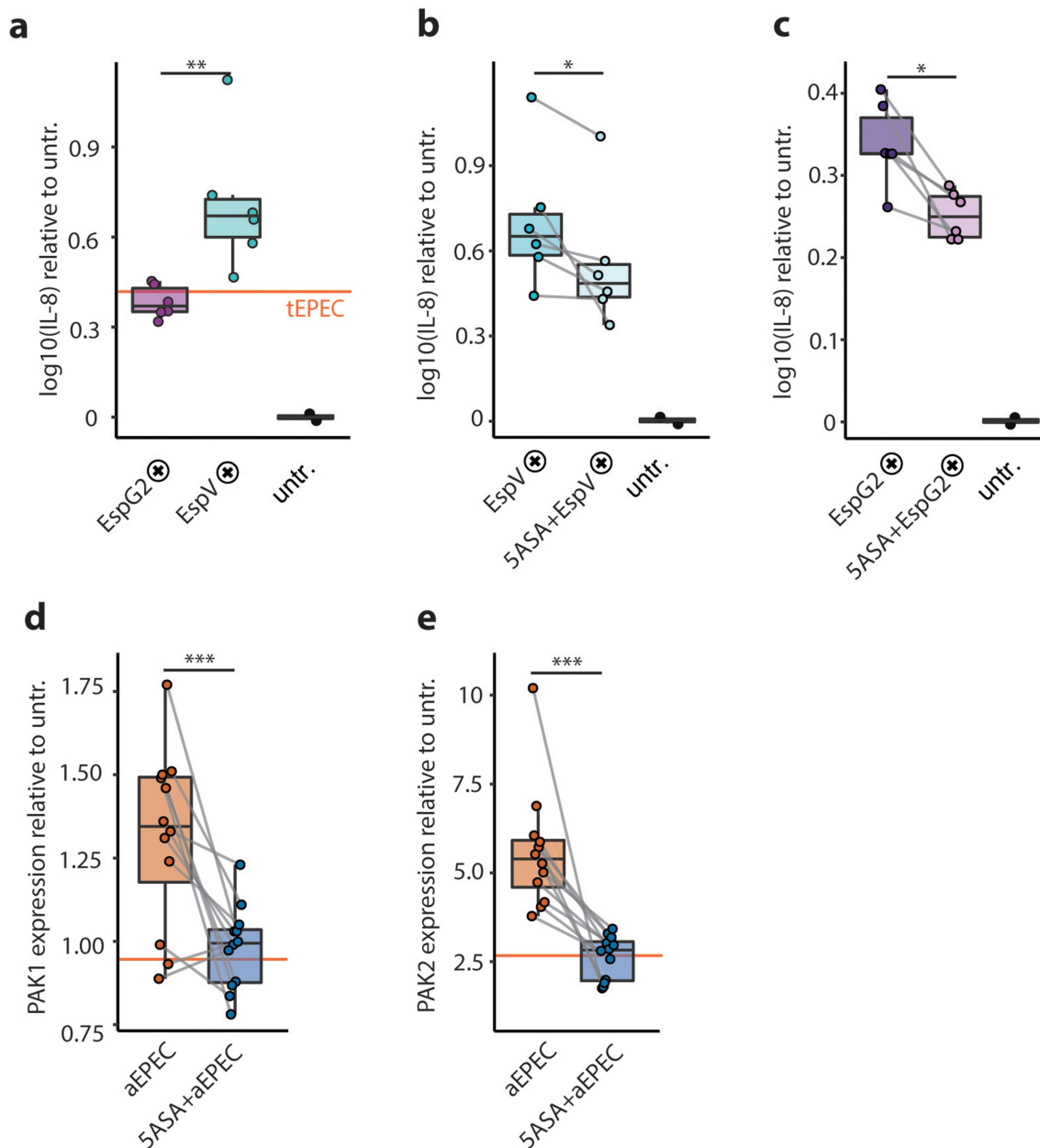
of aEPEC with human primary colon epithelial cells, EspG2-pos aEPEC had an IL-8 response comparable to the tEPEC-E2348/69 reference strain, while the median IL-8 secretion was doubled in response to EspV-pos strains (Figure 5a). The anti-inflammatory drug mesalamine (5-ASA) is the mainstay drug for mild-to-moderate UC. 5-ASA mediated inhibition of the master regulator PAK1 contributes to attenuation of multiple signaling pathways such as Wnt/ $\beta$ -catenin, ERK1/2, AKT1, mTOR, NF- $\kappa$ B and induction of cell cycle arrest.<sup>32,36,50</sup> EspG has been shown to activate PAKs, and PAK1 has recently been identified as an important driver of colitis in IBD in an integrated in vivo multiomics study.<sup>31,35</sup> Therefore, we investigated the effect of 5-ASA on aEPEC induced epithelial IL-8 secretion and PAK mRNA expression. Both IL-8 production and PAK1/2 expression were increased in aEPEC infected HCEC-1CT. 5-ASA treatment reduced the IL-8 response to EspG2-pos and EspV-pos aEPEC strains, together with a reduction of PAK1 and PAK2 mRNA expression (Figure 5b–e). 5-ASA treatment reduced the IL-8 response triggered by EspV-pos aEPEC to levels of EspG2-pos strains without 5-ASA (Figure 5b). With 5-ASA treatment, PAK expression of aEPEC infected cells was comparable to

tEPEC-E2348/69 infection (Figure 5d and e). Taken together, these findings support the hypothesis that 5-ASA can reduce the pro-inflammatory response elicited by EspV-pos aEPEC in UC patients.

## Discussion

The prevalence of aEPEC has surpassed tEPEC, with aEPEC being detected in 95–99% of The EPEC-positive samples in industrialized countries. In our study cohort, we detected aEPEC in ~10% and tEPEC in 0.4% of stool samples, which is in the range of other European studies.<sup>26,51</sup> The presence of aEPEC correlated with flares in UC disease activity as determined by fecal calprotectin, endoscopic and clinical-Mayo scores. Prevalence of aEPEC was similar in UC patients with active disease and health controls but halved in UC patients in remission. UC patients might represent a vulnerable population due to mutations in pathways involved in immune–microbiota interactions and intestinal barriers.<sup>1</sup> Comparing the microbiome composition of aEPEC-pos and aEPEC-neg UC samples, we detected reduced microbial diversity and an enrichment of ASVs belonging to taxa that also includes opportunistic pathogens *Dialister*, *Haemophilus*

and EHEC (red). (c) Pangenome of 57 AEEC investigated in this study, tree from alignment of 2719 orthologous proteins (left) with strains isolated from UC- (blue circle) and CD-patients (purple circle) labeled at the tip. Column depicts the presence of EspG2 (purple), EspV (blue), genome presence absence matrix (right) blue line below shows the percentage of isolates carrying a gene at each position and linearized (pan)genome, with genes displayed as rectangles above. Statistical analysis: (a) Fisher's exact test; \**p* ≤ .05.



**Figure 5.** 5-ASA dampens the aEPEC-induced pro-inflammatory epithelial response. (a) IL-8 secretion of human primary colon epithelial cells (HCEC-1CT) infected with EspG2- (purple) and EspV-positive (blue) aEPEC strains, normalized to untreated cells (black). (b) IL-8 secretion of primary colon epithelial cells infected with EspV-positive AEEC strains, without (blue) and with 5-ASA (light blue), normalized to untreated cells. (c) IL-8 secretion of primary colon epithelial cells infected with EspG2-positive AEEC strains, without (purple) and with 5-ASA (light purple), normalized to untreated cells. (d) PAK1 expression of primary colon epithelial cells infected with aEPEC, relative to untreated cells with (blue) and without (Orange) 5-ASA, tEPEC (Orange line). (e) PAK2 expression of primary colon epithelial cells infected with aEPEC, relative to untreated cells with (blue) and without (Orange) 5-ASA. (a-e) Values for the reference strain tEPEC-E2348/69 are visualized with an Orange line. (a-e) Data Points from the same strains are connected with a gray line. Statistical analysis: (a) unpaired t-test, (b-e) paired t-test, (a-c) Log<sub>10</sub> transformed y-axis; \* $p \leq .05$ , \*\* $p \leq .01$ , \*\*\* $p \leq .001$ .

and *Veillonella*. Correcting for calprotectin in differential abundance analysis pointed toward an independent influence of aEPEC on microbiome composition. *Haemophilus* is typical for the oral flora, and *Veillonella* is an important biofilm initiator.<sup>52</sup> Mucosal biofilms have recently been

shown to be correlated with dysbiosis and inflammation in IBD.<sup>44</sup> We found that aEPEC isolates from both UC and CD formed biofilms, especially under aerobic conditions. However, aEPEC isolated from healthy controls formed biofilms under anaerobic conditions. This suggests that aEPEC

isolated from IBD patients and healthy subjects differ in their adaptation to environmental parameters such as oxygen tension, which is likely a consequence of dysbiosis.

Compared to tEPEC, aEPEC persists longer in the intestine, which could be due to inhibition of epithelial apoptosis and its lack of localized adhesion.<sup>53</sup> We detected aEPEC positivity over several months in IBD patients together with cycles of reoccurrence. In our *in vitro* model, aEPEC were able to induce PAKs which are implicated in IBD pathogenesis and treatment with 5-ASA dampened PAK expression as well as IL-8 secretion. We have previously shown that PAK1 is overexpressed in IBD and is associated with increased cell survival.<sup>37</sup> Thus, activation of PAK signaling could contribute to intestinal persistence of aEPEC.

Isolated aEPEC strains from IBD patients showed *in vitro* phenotypes resembling diarrhea with increased paracellular flux after 4–6 hours of infection. Decreased TEER by aEPEC indicates compromised epithelial barrier which can be restored by 5-ASA.<sup>54</sup> Whether aEPEC infections precede and trigger IBD flares or aEPEC just thrives in an inflamed environment still remains to be determined. These results suggest that prolonged aEPEC infection could tip microbiota homeostasis and contribute to diarrhea and inflammation in UC patients. It is likely that aEPEC strains promote inflammation as a favorable ecological niche where they can outcompete commensals due to an abundance of virulence factors.

Human volunteer studies found that, contrary to tEPEC harboring *bfp*, the potential to cause diarrhea varies between different aEPEC strains and subjects.<sup>55–57</sup> tEPEC is known to compromise epithelial barrier and attenuate IL-8 secretion *in vitro*.<sup>49</sup> In contrast, aEPEC induced a stronger pro-inflammatory stimulus and less pronounced barrier defect which could be explained by the secretion of different effector proteins. aEPEC have a highly diverse virulence factor and effector protein repertoire due to evolution via repeated acquisition of LEE PAI variants and overall genetic plasticity.<sup>24</sup> It has been previously suggested that differences in the effector protein arsenal could explain the heterogeneous clinical phenotypes of aEPEC infection.<sup>25,26,51</sup> We showed that aEPEC from UC patients elicited a stronger epithelial IL-8 response

than aEPEC from CD, which behaved more like tEPEC. Virulence mechanisms and human target proteins of EspV are still elusive; however, their expression in yeast results in a dramatic increase in cell size and irreversible growth arrest.<sup>58</sup> Our analysis revealed that EspV-pos aEPEC were associated with UC and compared to EspG2-pos, had more virulence factors, including hemolysin E, adhesins, iron transporters and a T6SS combined with an elevated IL-8 response *in vitro*. Supporting the hypothesis of increased virulence potential, protein homology predicted an abundance of non-LEE T3SS effector proteins with unknown function that were enriched in EspV-pos aEPEC. The non-LEE effector protein EspG2 was associated with aEPEC from CD and could distinguish a phylogenetic aEPEC clade related to EPEC1. None of the aEPEC strains isolated from infectious diarrhea patients and just one of the strains from UC patients possessed EspG2. The preference of more virulent EspV-positive aEPEC for UC patients warrants further investigation. It might be explained by depleted mucin production, preexisting low-grade colonic inflammation or altered microbiome diversity.<sup>8,59</sup>

The modest sample size of our isolated strains is a limitation of this study. Possible sources of bias in the analysis are potential differences in sample storage between IBD and control cohort and age difference of healthy children vs. adult IBD patients for isolated strains. *E. coli* K-12 is known to disrupt the epithelial barrier and induces IL-8 secretion via TLR signaling *in vitro*.<sup>60–62</sup> Additional *in vivo* experiments, including mutant and non-aEPEC commensal strains isolated from controls, should be performed to establish EspG2 and EspV as *bona fide* genes for aEPEC virulence. Furthermore, detailed longitudinal studies could uncover the exact sequelae of aEPEC infection and the onset of inflammation in UC. aEPEC isolated in this study primarily belong to nonclassical EPEC serotypes that have not been recognized clinically. EspG2, EspV and the identified serotypes could serve as targets to distinguish aEPEC strains with different pro-inflammatory potentials in intestinal inflammation. *EspG2*, *espG* and *virA* from *Shigella flexneri* are structural homologies.<sup>63</sup> Future studies are vital to establish if EspG2 is just a marker of less co-occurring virulence factors in aEPEC genomes

or alters activation of pro-inflammatory signaling pathways such as PAK.

These results imply that EspV-pos aEPEC not only thrive in the niche of an inflamed GI environment but can also induce epithelial inflammation via induction of IL-8 secretion and PAK expression. The ability of aEPEC to generate disease depends on the susceptibility of the infected person, thus making aEPEC an opportunistic pathogen. Our findings suggest that UC patients with their disturbed immune-microbiota axis and less resilient microbiota might represent such a vulnerable population. Limiting the contact with aEPEC might thus contribute to secondary prevention in UC. In any case, stool samples should be screened for aEPEC in patients with flares of UC.

## Acknowledgments

We would like to acknowledge Anita Krnjic, Christina Gmainer, Marion Nehr, Helga Mock, and Sena Ecin for technical support in conducting the experiments.

## Disclosure statement

There is no competing interest.

## Funding

This study was supported by the Austrian Science Fund (P 32302) and the Vienna Science and Technology Fund (LS18-053; Austrian Science Fund (FWF)) [P 32302].

## References

- Graham DB, Xavier RJ. Pathway paradigms revealed from the genetics of inflammatory bowel disease. *Nature*. 2020;578:527–539. doi:10.1038/s41586-020-2025-2.
- Smillie CS, Biton M, Ordovas-Montanes J, Sullivan KM, Burgin G, Graham DB, Herbst RH, Rogel N, Slyper M, Waldman J, et al. Intra- and inter-cellular rewiring of the human colon during ulcerative colitis. *Cell*. 2019;178:714–730.e22. doi:10.1016/j.cell.2019.06.029.
- Evstatiev R, Cervenka A, Austerlitz T, Deim G, Baumgartner M, Beer A, Krnjic A, Gmainer C, Lang M, Frick A, et al. The food additive EDTA aggravates colitis and colon carcinogenesis in mouse models. *Sci Rep*. 2021;11:5188. doi:10.1038/s41598-021-84571-5.
- Ruiz PA, Morón B, Becker HM, Lang S, Atrott K, Spalinger MR, Scharl M, Wojtal KA, Fischbeck-Terhalle A, Frey-Wagner I, et al. Titanium dioxide

- nanoparticles exacerbate DSS-induced colitis: role of the NLRP3 inflammasome. *Gut*. 2017;66:1216–1224. doi:10.1136/gutjnl-2015-310297.
- Baumgart DC, Carding SR. Inflammatory bowel disease: cause and immunobiology. *Lancet*. 2007;369:1627–1640. doi:10.1016/S0140-6736(07)60750-8.
  - Maier L, Pruteanu M, Kuhn M, Zeller G, Telzerow A, Anderson EE, Brochado AR, Fernandez KC, Dose H, Mori H, et al. Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature*. 2018;555:623–628. doi:10.1038/nature25979.
  - Clooney AG, Eckenberger J, Laserna-Mendieta E, Sexton KA, Bernstein MT, Vagianos K, Sargent M, Ryan FJ, Moran C, Sheehan D, et al. Ranking microbiome variance in inflammatory bowel disease: a large longitudinal intercontinental study. *Gut*. 2021;70:499–510. doi:10.1136/gutjnl-2020-321106.
  - Lloyd-Price J, Arze C, Ananthakrishnan AN, Schirmer M, Avila-Pacheco J, Poon TW, Andrews E, Ajami NJ, Bonham KS, Brislawn CJ, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature*. 2019;569:655–662. doi:10.1038/s41586-019-1237-9.
  - Schirmer M, Denson L, Vlamakis H, Franzosa EA, Thomas S, Gotman NM, Rufo P, Baker SS, Sauer C, Markowitz J, et al. Compositional and temporal changes in the gut microbiome of pediatric ulcerative colitis patients are linked to disease course. *Cell Host Microbe*. 2018;24:600–610.e4. doi:10.1016/j.chom.2018.09.009.
  - Kotlowski R, Bernstein CN, Sepelhi S, Krause DO. High prevalence of *Escherichia coli* belonging to the B2+D phylogenetic group in inflammatory bowel disease. *Gut*. 2007;56:669–675. doi:10.1136/gut.2006.099796.
  - Petersen AM, Nielsen EM, Litrup E, Brynskov J, Mirsepasi H, Krogh KA. A phylogenetic group of *Escherichia coli* associated with active left-sided inflammatory bowel disease. *BMC Microbiol*. 2009;9:171. doi:10.1186/1471-2180-9-171.
  - Micenkova L, Frankovicova L, Jabornikova I, Bosak J, Dite P, Smarda J, Vrba M, Sevcikova A, Kmetova M, Smajs D, et al. *Escherichia coli* isolates from patients with inflammatory bowel disease: exPEC virulence- and colicin-determinants are more frequent compared to healthy controls. *Int J Med Microbiol*. 2018;308:498–504. doi:10.1016/j.ijmm.2018.04.008.
  - Desvaux M, Dalmasso G, Beyrouthy R, Barnich N, Delmas J, Bonnet R. Pathogenicity factors of genomic islands in intestinal and extraintestinal *Escherichia coli*. *Front Microbiol*. 2020;11:2065. doi:10.3389/fmicb.2020.02065.
  - Rasko DA, Rosovitz MJ, Myers GSA, Mongodin EF, Fricke WF, Gajer P, Crabtree J, Sebahia M, Thomson NR, Chaudhuri R, et al. The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *J Bacteriol*. 2008;190:6881–6893. doi:10.1128/JB.00619-08.

15. Baumgart M, Dogan B, Rishniw M, Weitzman G, Bosworth B, Yantiss R, Orsi RH, Wiedmann M, McDonough P, Kim SG, et al. Culture independent analysis of ileal mucosa reveals a selective increase in invasive *Escherichia coli* of novel phylogeny relative to depletion of Clostridiales in Crohn's disease involving the ileum. *ISME J.* 2007;1:403–418. doi:10.1038/ismej.2007.52.
16. Ryan P, Bennett MW, Aarons S, Lee G, Collins JK, O'Sullivan GC, O'Connell J, Shanahan F. PCR detection of mycobacterium paratuberculosis in Crohn's disease granulomas isolated by laser capture microdissection. *Gut.* 2002;51:665–670. doi:10.1136/gut.51.5.665.
17. Darfeuille-Michaud A, Boudeau J, Bulois P, Neut C, Glasser A-L, Barnich N, Bringer M-A, Swidsinski A, Beaugerie L, Colombel J-F, et al. High prevalence of adherent-invasive *Escherichia coli* associated with ileal mucosa in Crohn's disease. *Gastroenterology.* 2004;127(2):412–421. doi:10.1053/j.gastro.2004.04.061.
18. Sasaki M, Sitaraman SV, Babbitt BA, Gerner-Smidt P, Ribot EM, Garrett N, Alpern JA, Akyildiz A, Theiss AL, Nusrat A, et al. Invasive *Escherichia coli* are a feature of Crohn's disease. *Lab Invest.* 2007;87:1042–1054. doi:10.1038/labinvest.3700661.
19. Yang H, Mirsepasi-Lauridsen HC, Struve C, Allaire JM, Sivignon A, Vogl W, Bosman ES, Ma C, Fotovati A, Reid GS, et al. Ulcerative Colitis-associated *E. coli* pathobionts potentiate colitis in susceptible hosts. *Gut Microbes.* 2020;12:1847976. doi:10.1080/19490976.2020.1847976.
20. Limsrivilai J, Saleh ZM, Johnson LA, Stidham RW, Waljee AK, Govani SM, Gutermuth B, Brown AM, Briggs E, Rao K, et al. Prevalence and effect of intestinal infections detected by a PCR-based stool test in patients with inflammatory bowel disease. *Dig Dis Sci.* 2020;65:3287–3296. doi:10.1007/s10620-020-06071-2.
21. Axelrad JE, Joelson A, Green PHR, Lawlor G, Lichtiger S, Cadwell K, Lebowl B. Enteric infections are common in patients with flares of inflammatory bowel disease. *Am J Gastroenterol.* 2018;1–10. doi:10.1038/s41395-018-0211-8.
22. Elliott SJ, Wainwright LA, McDaniel TK, Jarvis KG, Deng Y, Lai L-C, McNamara BP, Sonnenberg MS, Kaper JB. The complete sequence of the locus of enterocyte effacement (LEE) from enteropathogenic *Escherichia coli* E2348/69. *Mol Microbiol.* 1998;28:1–4. doi:10.1046/j.1365-2958.1998.00783.x.
23. Kaper JB, Nataro JP, Mobley HL. Pathogenic *Escherichia coli*. *Nat Rev Microbiol.* 2004;2:123–140. doi:10.1038/nrmicro818.
24. Ingle DJ, Tauschek M, Edwards DJ, Hocking DM, Pickard DJ, Azzopardi KI, Amarasena T, Bennett-Wood V, Pearson JS, Tamboura B, et al. Evolution of atypical enteropathogenic *E. coli* by repeated acquisition of LEE pathogenicity island variants. *Nat Microbiol.* 2016;1. doi:10.1038/nmicrobiol.2015.10.
25. Hernandez RT, Elias WP, Vieira MAM, Gomes TAT. An overview of atypical enteropathogenic *Escherichia coli*. *FEMS Microbiol Lett.* 2009;297:137–149. doi:10.1111/j.1574-6968.2009.01664.x.
26. Hu J, Torres AG. Enteropathogenic *Escherichia coli*: foe or innocent bystander? *Clin Microbiol Infect.* 2015;21:729–734. doi:10.1016/j.cmi.2015.01.015.
27. Carlino MJ, Kralicek SE, Santiago SA, Sitaraman LM, Harrington AT, Hecht GA. Quantitative analysis and virulence phenotypes of atypical enteropathogenic *Escherichia coli* (EPEC) acquired from diarrheal stool samples from a Midwest US hospital. *Gut Microbes.* 2020;12:1–21. doi:10.1080/19490976.2020.1824562.
28. Sharma R. Balance of bacterial pro- and anti-inflammatory mediators dictates net effect of enteropathogenic *Escherichia coli* on intestinal epithelial cells. *Ajp.* 2006;290:G685–G694.
29. Slater SL, Frankel G. Advances and challenges in studying type III secretion effectors of attaching and effacing pathogens. *Front Cell Infect Microbiol.* 2020;10:337. doi:10.3389/fcimb.2020.00337.
30. Dean P, Kenny B. The effector repertoire of enteropathogenic *E. coli*: ganging up on the host cell. *Curr Opin Microbiol.* 2009;12:101–109. doi:10.1016/j.mib.2008.11.006.
31. Selyunin AS, Sutton SE, Weigele BA, Reddick LE, Orchard RC, Bresson SM, Tomchick DR, Alto NM. The assembly of a GTPase–kinase signalling complex by a bacterial catalytic scaffold. *Nature.* 2011;469:107–111. doi:10.1038/nature09593.
32. Dammann K, Khare V, Gasche C. Tracing PAKs from GI inflammation to cancer. *Gut.* 2014;63:1173–1184. doi:10.1136/gutjnl-2014-306768.
33. Singh V, Davidson A, Hume PJ, Koronakis V. Pathogenic *Escherichia coli* hijacks GTPase-activated p21-activated kinase for actin pedestal formation. *mBio.* 2019;10. doi:10.1128/mbio.01876-19.
34. Sun H, Kamanova J, Lara-Tejero M, Galán JE. Salmonella stimulates pro-inflammatory signalling through p21-activated kinases bypassing innate immune receptors. *Nat Microbiol.* 2018;3:1122–1130. doi:10.1038/s41564-018-0246-z.
35. Lyons J, Brubaker DK, Ghazi PC, Baldwin KR, Edwards A, Boukhali M, Strasser SD, Suarez-Lopez L, Lin Y-J, Yajnik V, et al. Integrated in vivo multiomics analysis identifies p21-activated kinase signaling as a driver of colitis. *Sci Signal.* 2018;11. doi:10.1126/scisignal.aan3580.
36. Dammann K, Khare V, Lang M, Claudel T, Harpain F, Granofszky N, Evstatiev R, Williams JM, Pritchard DM, Watson A, et al. PAK1 modulates a PPAR $\gamma$ /NF- $\kappa$ B cascade in intestinal inflammation. *Biochimica et biophysica acta.* 2015;1853:2349–2360. doi:10.1016/j.bbamcr.2015.05.031.
37. Khare V, Dammann K, Asboth M, Krnjic A, Jambrich M, Gasche C. Overexpression of PAK1 promotes cell survival in inflammatory bowel diseases and colitis-associated cancer. *Inflamm Bowel Dis.* 2015;21:287–296. doi:10.1097/MIB.0000000000000281.

38. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods*. 2016;13:581–583. doi:10.1038/nmeth.3869.
39. Lagkouvardos I, Fischer S, Kumar N, Clavel T. Rhea: a transparent and modular R pipeline for microbial profiling based on 16S rRNA gene amplicons. *PeerJ*. 2017;5:e2836. doi:10.7717/peerj.2836.
40. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550. doi:10.1186/s13059-014-0550-8.
41. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol*. 2020;37:1530–1534. doi:10.1093/molbev/msaa015.
42. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J, et al. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics*. 2015;31:3691–3693. doi:10.1093/bioinformatics/btv421.
43. Suchsmita A, Jha A. IDDF2019-ABS-0129 optimal cut-off value of fecal calprotectin for the evaluation of inflammatory bowel disease: an unsolved issue? *Gut*. 2019;68:A85–A86.
44. Baumgartner M, Lang M, Holley H, Crepez D, Hausmann B, Pjevac P, Moser D, Haller F, Hof F, Beer A, et al. Mucosal biofilms are an endoscopic feature of irritable bowel syndrome and ulcerative colitis. *Gastroenterology*. 2021;161:1245–1256.e20. doi:10.1053/j.gastro.2021.06.024.
45. Fitzhenry RJ, Stevens MP, Jenkins C, Wallis TS, Heuschkel R, Murch S, Thomson M, Frankel G, Phillips AD. Human intestinal tissue tropism of intimin epsilon O103 *Escherichia coli*. *FEMS Microbiol Lett*. 2003;218:311–316. doi:10.1016/S0378-1097(02)01182-5.
46. Mundy R, Schüller S, Girard F, Fairbrother JM, Phillips AD, Frankel G. Functional studies of intimin in vivo and ex vivo: implications for host specificity and tissue tropism. *Microbiology*. 2007;153:959–967. doi:10.1099/mic.0.2006/003467-0.
47. Reid SD, Herbelin CJ, Bumbaugh AC, Selander RK, Whittam TS. Parallel evolution of virulence in pathogenic *Escherichia coli*. *Nature*. 2000;406:64–67. doi:10.1038/35017546.
48. Smollett K, Shaw RK, Garmendia J, Knutton S, Frankel G. Function and distribution of EspG2, a type III secretion system effector of enteropathogenic *Escherichia coli*. *Microbes Infect*. 2006;8:2220–2227. doi:10.1016/j.micinf.2006.04.004.
49. Ruchaud-Sparagano MH, Maresca M, Kenny B. Enteropathogenic *Escherichia coli* (EPEC) inactivate innate immune responses prior to compromising epithelial barrier function. *Cell Microbiol*. 2007;9:1909–1921. doi:10.1111/j.1462-5822.2007.00923.x.
50. Frick A, Khare V, Jimenez K, Dammann K, Lang M, Krnjic A, Gmainer C, Baumgartner M, Mesteri I, Gasche C, et al. A novel PAK1–Notch1 axis regulates crypt homeostasis in intestinal inflammation. *Cell Mol Gastroenterol Hepatol*. 2021;11:892–907.e1. doi:10.1016/j.jcmgh.2020.11.001.
51. Ochoa TJ, Barletta F, Contreras C, Mercado E. New insights into the epidemiology of enteropathogenic *Escherichia coli* infection. *Trans R Soc Trop Med Hyg*. 2008;102:852–856. doi:10.1016/j.trstmh.2008.03.017.
52. Mashima I, Nakazawa F. A review on the characterization of a novel oral Veillonella species, *V. Tobetsuensis*, and its role in oral biofilm formation. *J Oral Biosci*. 2013;55:184–190. doi:10.1016/j.job.2013.07.002.
53. Nguyen RN, Taylor LS, Tauschek M, Robins-Browne RM. Atypical enteropathogenic *Escherichia coli* infection and prolonged diarrhea in children. *Emerg Infect Dis*. 2006;12:597–603. doi:10.3201/eid1204.051112.
54. Khare V, Krnjic A, Frick A, Gmainer C, Asboth M, Jimenez K, Lang M, Baumgartner M, Evstatiev R, Gasche C, et al. Mesalamine and azathioprine modulate junctional complexes and restore epithelial barrier function in intestinal inflammation. *Sci Rep*. 2019;9:2842. doi:10.1038/s41598-019-39401-0.
55. Levine MM, Nalin D, Hornick R, Bergquist E, Waterman D, Young C, Sotman S, Rowe B. *Escherichia coli* strains that cause diarrhoea but do not produce heat-labile or heat-stable enterotoxins and are non-invasive. *Lancet*. 1978;1:1119–1122. doi:10.1016/S0140-6736(78)90299-4.
56. Bieber D, Ramer SW, Wu C-Y, Murray WJ, Tobe T, Fernandez R, Schoolnik GK. Type IV Pili, transient bacterial aggregates, and virulence of enteropathogenic *Escherichia coli*. *Science*. 1998;280:2114–2118. doi:10.1126/science.280.5372.2114.
57. Levine MM, Nataro JP, Karch H, Baldini MM, Kaper JB, Black RE, Clements ML, O'Brien AD. The diarrheal response of humans to some classic serotypes of enteropathogenic *Escherichia coli* is dependent on a plasmid encoding an enteroadhesiveness factor. *J Infect Dis*. 1985;152:550–559. doi:10.1093/infdis/152.3.550.
58. Arbeloa A, Oates CV, Marchès O, Hartland EL, Frankel G, Baumler AJ. Enteropathogenic and enterohemorrhagic *Escherichia coli* type III secretion effector espv induces radical morphological changes in eukaryotic cells. *Infect Immun*. 2011;79:1067–1076. doi:10.1128/IAI.01003-10.
59. Alipour M, Zaidi D, Valcheva R, Jovel J, Martínez I, Sergi C, Walter J, Mason AL, Wong GKS, Dieleman LA, et al. Mucosal barrier depletion and loss of bacterial diversity are primary abnormalities in paediatric ulcerative colitis. *J Crohns Colitis*. 2016;10:462–471. doi:10.1093/ecco-jcc/jjv223.
60. Bambou J-C, Giraud A, Menard S, Begue B, Rakotobe S, Heyman M, Taddei F, Cerf-Bensussan N, Gaboriau-Routhiau V. In vitro and ex vivo



- activation of the TLR5 signaling pathway in intestinal epithelial cells by a commensal *Escherichia coli* strain. *J Biol Chem.* 2004;279:42984–42992. doi:10.1074/jbc.M405410200.
61. Bhat MI, Sowmya K, Kapila S, Kapila R. *Escherichia coli* K12: an evolving opportunistic commensal gut microbe distorts barrier integrity in human intestinal cells. *Microb Pathog.* 2019;133:103545. doi:10.1016/j.micpath.2019.103545.
  62. Bereswill S, Fischer A, Dunay IR, Kühl AA, Göbel UB, Liesenfeld O, Heimesaat MM. Pro-inflammatory potential of *Escherichia coli* strains K12 and Nissle 1917 in a murine model of acute ileitis. *Eur J Microbiol Immunol.* 2013;3:126–134. doi:10.1556/EuJMI.3.2013.2.6.
  63. Davis J, Wang J, Tropea JE, Zhang D, Dauter Z, Waugh DS, Wlodawer A. Novel fold of VirA, a type III secretion system effector protein from *Shigella flexneri*. *Protein Sci.* 2008;17:2167–2173. doi:10.1110/ps.037978.108.