

Characterization of clumpy adhesion of *Escherichia coli* to human cells and associated factors influencing antibiotic sensitivity

Muhammad Moman Khan,¹ Katarzyna Sidorczuk,² Juliane Becker,¹ Adrianna Aleksandrowicz,³ Karolina Baraniewicz,^{3,4} Christina Ludwig,⁵ Aamir Ali,⁶ Robert A. Kingsley,⁷ Peter Schierack,¹ Rafał Kolenda^{3,7}

AUTHOR AFFILIATIONS See affiliation list on p. 17.

ABSTRACT *Escherichia coli* intestinal infection pathotypes are characterized by distinct adhesion patterns, including the recently described clumpy adhesion phenotype. Here, we identify and characterize the genetic factors contributing to the clumpy adhesion of *E. coli* strain 4972. In this strain, the transcriptome and proteome of adhered bacteria were found to be distinct from planktonic bacteria in the supernatant. A total of 622 genes in the transcriptome were differentially expressed in bacteria present in clumps relative to the planktonic bacteria. Seven genes targeted for disruption had variable distribution in different pathotypes and nonpathogenic *E. coli*, with the *pilV* and *spnT* genes being the least frequent or absent from most groups. Deletion (Δ) of five differentially expressed genes, *flgH*, *ffp*, *pilV*, *spnT*, and *yggT*, affected motility, adhesion, or antibiotic stress. Δ *flgH* exhibited 80% decrease and Δ *yggT* depicted 184% increase in adhesion, and upon complementation, adhesion was significantly reduced to 13%. Δ *flgH* lost motility and was regenerated when complemented, whereas Δ *ffp* had significantly increased motility, and reintroduction of the same gene reduced it to the wild-type level. The clumps produced by Δ *ffp* and Δ *spnT* were more resistant and protected the bacteria, with Δ *spnT* showing the best clump formation in terms of ampicillin stress protection. Δ *yggT* had the lowest tolerance to gentamicin, where the antibiotic stress completely eliminated the bacteria. Overall, we were able to investigate the influence of clump formation on cell surface adhesion and antimicrobial tolerance, with the contribution of several factors crucial to clump formation on susceptibility to the selected antibiotics.

IMPORTANCE The study explores a biofilm-like clumpy adhesion phenotype in *Escherichia coli*, along with various factors and implications for antibiotic susceptibility. The phenotype permitted the bacteria to survive the onslaught of high antibiotic concentrations. Profiles of the transcriptome and proteome allowed the differentiation between adhered bacteria in clumps and planktonic bacteria in the supernatant. The deletion mutants of genes differentially expressed between adhered and planktonic bacteria, i.e., *flgH*, *ffp*, *pilV*, *spnT*, and *yggT*, and respective complementations in *trans* cemented their roles in multiple capacities. *ffp*, an uncharacterized gene, is involved in motility and resistance to ampicillin in a clumpy state. The work also affirms for the first time the role of the *yggT* gene in adhesion and its involvement in susceptibility against another aminoglycoside antibiotic, i.e., gentamicin. Overall, the study contributes to the mechanisms of biofilm-like adhesion phenotype and understanding of the antimicrobial therapy failures and infections of *E. coli*.

KEYWORDS *Escherichia coli*, clumpy adhesion, motility, antibiotic tolerance, stress, adhesion

Escherichia coli is present in the gut microbiota of animals soon after birth, contributing to colonization resistance to pathogens, but it may also cause diseases that

Editor Ryan Rego, Institute of Parasitology, Biology Centre, ASCR, Ceske Budejovice, Czechia

Address correspondence to Rafał Kolenda, rafal.kolenda@upwr.edu.pl.

The authors declare no conflict of interest.

Received 24 June 2023

Accepted 12 March 2024

Published 26 March 2024

Copyright © 2024 Khan et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

have the potential to move across and affect multiple hosts and be fatal (1). Disseminated infections may require treatment with antibiotics and although *E. coli* is inherently sensitive to almost all classes of antibiotics, frequent acquisition of resistance genes through horizontal gene transfer complicates treatment and contributes to resistance in other pathogens (2). *E. coli* is the leading cause of death from bacterial infections, attributable to and associated with antimicrobial resistance, totaling over 800,000 per year (3). *E. coli* is a priority pathogen included in the list of multidrug-resistant bacteria issued by the World Health Organization, which needs to be addressed to counter this global public health challenge (4).

Colonization by *E. coli* pathotypes is initiated by contact with the host gut mucosal epithelium, and adherence is important in overcoming mechanical removal and clearance of pathogens from the gut (5). Extracellular appendages and adhesins play a critical role in the specific binding of the bacterial cell to receptors on the host cell (6). Adhesion also facilitates the deployment of virulence factors that may include the transfer of effector proteins into host enterocytes through type III secretion systems. Diverse mechanisms of colonization and infection result in a variety of adhesion patterns that are associated with specific *E. coli* pathotypes. Enteroaggregative *E. coli* (EAEC) display an aggregative adherence pattern or localized adhesion due to the autoagglutination of bacterial cells shown by the stacked-brick configuration, which occurs on the host cell surface usually in the presence of D-mannose. Similarly, enteropathogenic *E. coli* (EPEC) also portrayed an aggregative adherence pattern mediated by specific fimbriae encoded on plasmids. A specialized adherence mechanism known as the attaching and effacing lesion involves intimate attachment to the host surface, followed by effacement of microvilli and cytoskeletal rearrangements. Diffusely adherent *E. coli* (DAEC) are characterized by scattered distribution over the host cell surface, and the adhesion pattern is referred to as diffuse adherence (7, 8). The importance of adhesion patterns in pathogenesis is not well understood, but they may also affect the sensitivity to antimicrobial agents or other environmental stressors present in the intestine due to decreased penetration of the stressor, in contrast to diffusely distributed single bacteria (9).

In a study of 282 *E. coli* strains isolated from diverse host species, four distinct adhesion patterns on a variety of cell types and tissues were observed: (i) diffusely distributed single bacteria; (ii) microcolonies; (iii) chains; and (iv) clumps (10). Adhesion as clumps was rare, only observed for two *E. coli* strains, and more pronounced on the human urinary bladder cell line 5637 (10). Clumps of bacteria were large, three-dimensional structures similar to those of biofilms and composed of several thousand bacteria that formed after only 4 h of incubation. Clumps constituted larger and more complex structures covering the cell line extensively and were resistant to repeated washing, whereas microcolonies formed simpler aggregates of clustered bacteria when visualized under the fluorescent microscope. Little is known about the bacterial factors required for clump formation and its impact on resistance to stressors. We, therefore, tested the hypothesis that bacteria present in *E. coli* strain 4972 clumps exhibit distinct gene expression profiles and that some genes with altered expression are involved in clumpy adhesion. We also investigated the contribution of *E. coli* strain 4972-encoded factors that affected clump formation on the susceptibility to the antibiotics.

MATERIALS AND METHODS

Bacterial isolates and plasmids

The strains and plasmids used in this study are listed in Table S1.

Next-generation sequencing and genome analysis

Genomic DNA of *E. coli* strain 4972 was extracted using Wizard Genomic DNA Purification Kit (Promega Corp., Madison, USA) from 5 mL overnight culture as per the manufacturer's

instructions. The extracted nucleic acid was sequenced (Illumina), assembled with Shovill (11), and annotated using PROKKA (12) software. The whole genome sequence of strain 4972 is freely available from the GenBank database under accession number [GCA_018429805.1](https://www.ncbi.nlm.nih.gov/nuccore/GCA_018429805.1) (13). To compare the phylogenetic relationship of strain 4972 in the context of previously published *E. coli* genomes, a similar approach described previously was selected (14). Genome sequences of *E. coli* genomes were downloaded from NCBI GenBank using accession numbers (Table S2), and SNPs in the core genome alignment were identified using snippy software, with reference to *E. coli* K-12 MG1655 (GenBank accession no.: [CP005930](https://www.ncbi.nlm.nih.gov/nuccore/CP005930)). A phylogenetic tree was constructed using RAxML (15). Phylogroup and pathotype information for reference strains (Table S2) was described previously (14) and annotated on the tree using ggtree R package (16).

Adhesion assays

The cell line of a urinary bladder carcinoma—5637, was cultivated in RPMI medium with the addition of 10% fetal bovine serum, 2 mM L-glutamine, 1 mM sodium pyruvate, and 1% penicillin/streptomycin. For the assay, a medium without antibiotics was used (later referred to as infection medium). Cells were grown at 37°C, 95% humidity, and 5% CO₂ and passaged when the flask was of 80% ± 10% confluent. To visualize and semi-quantify the bacterial clumps, an adhesion assay was conducted in a 96-well format where the wild-type strain 4972, 4972 transformed with pFPV25.1Kan (encoding GFP), and deletion mutants were allowed to adhere to the monolayer of 5637 cells for 4 h, stained with propidium iodide (PI) or DAPI (4',6-diamidino-2-phenylindole), and visualized using automated fluorescence microscopy-based Aklides system according to the already published methodology (10).

As it was not possible to fully quantify the clumpy adhesion phenotype of the mutants with respect to the wild type, quantification of colony-forming units (CFU) was done. For this, 5637 cells were seeded onto 24-well plates, and the assay was conducted when a confluent monolayer had formed. Prior to the addition of bacterial strains, monolayers of 5637 were washed with PBS and then 500 µL of infection medium was pipetted in each well. The day before reaching the monolayer, chosen bacteria were inoculated in 1 mL of LB and incubated at 180 rpm at 37°C for 16 h. The next day, OD₆₀₀ of the cultures was measured and diluted to OD₆₀₀ = 0.05 in 5 mL LB medium and grown at 220 rpm, 37°C to OD₆₀₀ = 2.0. If a mutant with a pBAD33 expression vector was used, L-arabinose was added with a final concentration of 0.2% after reaching OD₆₀₀ = 0.5, and bacteria were allowed to grow for another 2 h. After that, 1 mL of bacterial culture was washed with PBS, resuspended in 1 mL of infection medium, and bacterial OD₆₀₀ was determined. For each strain, 300 µL of infection medium containing 1.2 × 10⁵ bacteria was added to each of the four well replicates. The plates were incubated at 37°C for 4 h, and each well was washed three times with 1 mL of PBS. Next, 1 mL of PBS with 1% Triton X-100 was added per well, the plate was incubated for 5 min at room temperature, and finally, the contents of the well were scraped by pipetting in each well and mixed thoroughly. A volume of 200 µL from each well was then transferred into 96-well plates and a dilution series was generated. On LB agar plates, 20 µL spots of the dilutions were pipetted in triplicates for each well. The plates were placed in the incubator for overnight growth at 37°C, and bacterial colonies were counted the next morning.

Bacteria within biofilm-like clumps were stressed using a modification of the adhesion assay in 24-well plates. Initially, dose-ranging experiments for ampicillin and gentamicin were carried out by a gradual increase in concentration in the infection medium after the clump formation of the wild-type strain. For ampicillin, the concentration tested for stress ranged from 0 to 16,000 µg/mL, whereas for gentamicin 0–16 µg/mL were tested. The concentration for exposure to each antibiotic, which resulted in a decrease in more than half of the viable bacteria, was selected. After the 4-h incubation and clump formation, the infection media were aspirated, and a fresh infection medium containing the antibiotics was added. The plate was incubated for another 1 h at 37°C. The rest of the procedure was identical to that for the adhesion assay.

Biofilm assay

The strain 4972 was tested for biofilm formation using VideoScan in a variety of media: LB, LB supplemented with glucose (LB+), tryptic soy broth (TSB), TSB supplemented with glucose (TSB+), brain heart infusion (BHI) media, BHI supplemented with glucose (BHI+), M63 minimal media, M63 minimal media supplemented with glucose (M63+), RPMI cell culture media, and DMEM cell culture media using the already established protocol (17).

RNA isolation from *E. coli* strain 4972 and sequencing

Adhesion assay on three different cell lines, i.e., porcine kidney (PK15), human intestinal (Caco-2), and urinary bladder carcinoma 5637, was carried out, where strain 4972 was allowed to adhere to these cell lines for 4 h with their respective mediums as detailed in Table S3. The procedure was the same for each cell line, whereby unattached bacteria in the supernatant were collected, the adhered bacterial cells were carefully removed by pipetting, and the resultant fluid from both was subjected to RNA isolation by the RNAsnap method (18). RNA was sequenced with 100-bp paired-end read (Illumina HiSeq2000) at LGC's Agowa Genomics, Berlin, Germany. The library for sequencing was prepared with a NEB Ultra II RNA custom kit (NEB). Sequencing quality was assessed on the basis of average base quality, GC content, and adapter contamination as described previously (19). RNA sequences of adhered and unattached bacterial strain 4972 are available in the NCBI BioProject database under accession number PRJNA868552. Both alignment and feature counting were performed using the R package Rsubread (20). Differentially expressed genes (DEGs) were identified with a 0.05 false discovery rate (FDR) threshold using the exact test of edgeR (21, 22). Leading logFC values were used to generate an MDS plot using ggplot2 (23).

Sequences of DEGs unannotated by PROKKA ($n = 148$) along with annotated target genes were analyzed by NCBI blastp, Protein Homology/analogy Recognition Engine (Phyre2) (24), bacterial localization prediction tool (PSORTb 3.0) (25), transmembrane topology and signal peptide predictor (Phobius) (26), and Prediction of transmembrane helices in proteins (TMHMM Server v.2.0.) (27). DEGs were functionally annotated using eggnoG-mapper (28) and assigned functions/categories according to the database of Clusters of Orthologous Genes (COGs) (29, 30). Next, the gene set enrichment analysis (GSEA) function from the clusterProfiler package was used to perform a gene set enrichment analysis for COG functions/categories (31). COG functions/categories were considered significantly upregulated/downregulated if Benjamini-Hochberg corrected P value was equal to or lower than 0.05.

Bacterial gene deletion mutant generation

The Datsenko-Wanner recombineering method was applied with minor modifications for gene deletion of *flgH* and *spy* genes (32). For the deletion of *fimH*, *ffp*, *yggT*, *pilV*, and *spnT*, a method by Derous and Deboeck (33) was utilized. Positive clones were streaked on agar plates with appropriate antibiotics, and colony PCRs were performed to check and confirm successful gene disruption. The primers used are all listed in Table S4.

The whole genome of the gene deletion mutants was extracted by Qiagen Cube using the Qiagen DNA extraction kit and was sent to Novogene, UK for whole-genome sequencing with Illumina NovaSeq 6000 (Illumina). For single nucleotide variation search, the reads obtained for each isolate were mapped against the strain 4972 genome using bwa and called with bcftools (34). In the case of $\Delta spnT$, one mutation before the start codon of the gene *nemR* was detected and was used for further analysis as it did not exhibit any irregularities in the growth of the strain. Sequences of all the deletion mutants for this study are available in the NCBI BioProject database under accession number PRJNA880721. The antibiotic resistance cassette in deletion mutants was removed using the pCP20 plasmid as described previously (35).

Cloning of genes from strain 4972 into pBAD33 plasmid and complementation in mutants

For complementation of deletion mutants, only mutants exhibiting altered adhesion in initial experiments, i.e., *flgH*, *ffp*, *yggT*, *spnT*, and *pilV*, were cloned from the wild-type strain 4972 into expression vector pBAD33 as described previously (36). Five mutants (i.e., $\Delta flgH$, Δffp , $\Delta yggT$, $\Delta spnT$, and $\Delta pilV$) were transformed with a plasmid carrying corresponding complementing gene as well as pBAD33 vector plasmid alone (without insert) as a negative control (35).

Minimum inhibitory concentration assay

Minimum inhibitory concentration (MIC) assay of the wild-type strain 4972 and gene mutants, i.e., Δffp , $\Delta yggT$, $\Delta spnT$, and $\Delta pilV$, was carried out on a range of antibiotic concentrations of ampicillin (0–32 $\mu\text{g}/\text{mL}$) and gentamicin (0–8 $\mu\text{g}/\text{mL}$). For this, 1 mL of bacterial strain culture was incubated for 16 h at 37°C and 180 rpm. The OD₆₀₀ was determined, and bacteria were diluted in the infection media to get a total volume of 1 mL containing 5×10^6 bacteria. This was then aliquoted to 200 μL per well into a 96-well Nunclon plate and incubated at 37°C. After the 24 h, OD₆₀₀ was measured for each well with the Tecan microplate reader (Tecan, Austria). Three technical and biological repetitions were performed for each strain. The lowest concentration inhibiting growth, i.e., MIC, was defined as OD₆₀₀, where no growth was observed, mirroring the optical density of the negative control.

Motility assays

Motility assays of the wild-type strain 4972, deletion mutants, complemented mutants, and mutants with empty plasmid were conducted according to the already published protocol (17, 36).

Prevalence of target genes in different pathotypes

Genes targeted for deletion were analyzed for prevalence in various *E. coli* pathotypes using Abricate and *E. coli* genomes downloaded from the RefSeq database (accessed on 26 October 2021) (37). The 25,436 downloaded genomes were filtered to exclude those of low quality (coverage below 50, third generation, or unknown sequencing platform, more than 400 contigs) and then pathotyped by running a BLAST search of selected genes associated with various pathotypes against the database of genomes. Genes with *E*-value $\leq 10^{-100}$ were considered as present (13). Decisions used for pathotyping and the numbers of genomes assigned to each pathotype are shown in Table S5.

Global proteome analysis of *E. coli* clumps

For the proteomic analysis of clumps and planktonic bacteria, the adhesion assay of the wild-type strain 4972 on cell line 5637 was carried out as described above. The planktonic bacteria in the supernatant were removed from each well of the 24-well plate and pooled using a pipette. The adhered bacterial cells were carefully detached by aspirating ice-cold 1× PBS in each well of the 24-well plate using a pipette. Bacterial cells were washed with ice-cold 1× PBS and finally resuspended in 1 mL of 1× PBS. Bacterial cells were resuspended in freshly prepared 8 M urea lysis buffer (8 M Urea, 5 mM EDTA, 100 mM NH₄HCO₃, and 1 mM DTT, pH 8.0) and lysed by sterile bead (500 mg glass beads) beating for 3 min with a frequency of 30 s⁻¹ by TissueLyser II (Qiagen). The experiment and protein extraction were carried out five times independently, i.e., five biological replicates.

Protein extracts were sent to the Bavarian Center for Biomolecular Mass Spectrometry (BayBioMS) for proteomic analysis. Per sample, 15 μg of protein extract was reduced (10 mM DTT), carbamidomethylated (55 mM CAA), and digested with trypsin (0.3 μg Trypsin Gold, Promega). Generated peptides (0.2 μg per injection) were analyzed on a

Dionex Ultimate 3000 RSLCnano system coupled to a Q-Exactive HF-X mass spectrometer (ThermoFisher Scientific) using a 50 min linear gradient from 4% to 32% of solvent B [0.1% (vol/vol) formic acid and 5% (vol/vol) dimethyl sulfoxide (DMSO) in acetonitrile] at 300 nL/min flow rate. nanoLC solvent A was 0.1% (vol/vol) formic acid and 5% (vol/vol) DMSO in HPLC-grade water. The Q-Exactive HF-X mass spectrometer was operated in data-dependent acquisition and positive ionization modes. Peptide fragmentation was performed using higher energy collision-induced dissociation and a normalized collision energy of 26%.

Peptide identification and quantification were done using the software MaxQuant (version 1.6.3.4) (38). MS2 spectra were searched against the proteins from PROKKA annotated strain 4972 whole genome, supplemented with common contaminants (built-in option in MaxQuant). The results were adjusted to 1% FDR on peptide spectrum match level and protein level, employing a target-decoy approach using reversed protein sequences. The proteins differentially expressed between adhered and planktonic bacterial cells were identified using the label-free quantification algorithm provided by MaxQuant (LFQ) (39) and further analyzed with the limma R package. Briefly, the data provided were processed by first filtering out the peptides, which equaled 1. Rows with missing intensity values for at least four samples were filtered, and LFQ intensity values were logarithmized (log2). The missing values were imputed using the MinDet method from imputeLCMD. A linear model was fitted to the data using moderated *t*-statistics, and log odds of differential expression were calculated by empirical Bayes moderation of the standard errors toward a global value using the limma package. *P*-values were adjusted for multiple testing using the Benjamini-Hochberg method. DEGs were identified with a 0.05 false discovery rate threshold and logFC values below -0.5 or above 0.5 . The proteins were functionally annotated with the use of an eggNOG-mapper and used for gene set enrichment analysis, as mentioned previously in section "Biofilm assay" (28).

Statistical analyses

Statistical analysis was carried out using the R statistical software (40). Gene prevalence was compared using the Chi-squared test of independence implemented in R. Figures were generated with the use of the ggplot2 package implemented in R software (41). Wilcoxon rank-sum test was used to analyze the differences, and Benjamini-Hochberg (*P*.BH) lower than 0.05 was considered significant.

RESULTS

Clump formation by *E. coli* 4972 is specific to epithelial cells

We previously defined a new adhesion phenotype that was demonstrated by two *E. coli* strains, one isolated from human, i.e., 4972, was used in this study, and the other from squirrel. The adhesion pattern exhibited a biofilm-like structure, referred to as clumps, within 4 h on human urinary bladder epithelial cell line 5637 (10). Clumps were less pronounced on PK15 and Caco-2 cell lines. *E. coli* strain 4972 formed a weak biofilm on plastic wells (Fig. S1) but an extensive clumpy structure on the 5637 bladder cell line (Fig. 1).

Strain 4972 shares phylogenomic placement with diverse *E. coli* pathotypes

The phylogenetic relationship of *E. coli* strain 4972 was investigated in the context of a diverse collection of *E. coli* reference strains from multiple phylogroups and pathotypes, including commensal strains. A phylogenetic tree was constructed based on SNPs in the shared genome with reference to *E. coli* K-12 MG1655 (Fig. 2). *E. coli* strain 4972 was present within a clade corresponding to phylogroup A (14) containing strains isolated from human fecal samples with diverse lifestyles, including commensal, enteroinvasive *E. coli* (EIEC), EAEC, enterohemorrhagic *E. coli* (EHEC), and enterotoxigenic *E. coli* (ETEC) pathotypes.

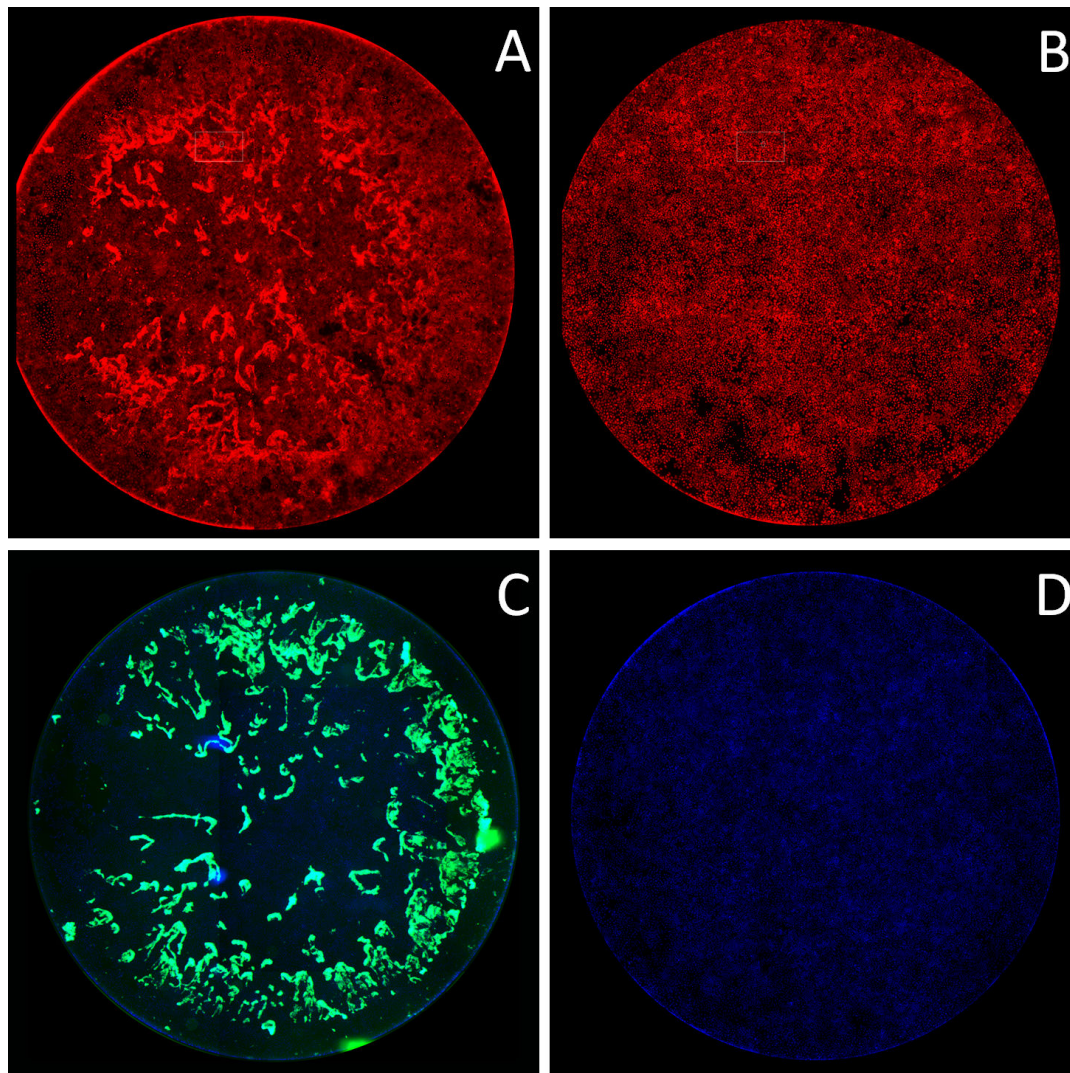


FIG 1 Clump formation by strain 4972. (A) PI-stained wild-type *E. coli* strain 4972, (B) PI-stained 5637 cell line only as a negative control, (C) GFP-expressing 4972 + pFPV25.1 Kan and DAPI-stained nuclei of 5637 cell line, and (D) DAPI-stained nuclei of 5637 cell line only as a negative control, after 4-h incubation of adhesion assay and visualized by Aklides system.

Transcriptome reveals differences among the subpopulation of sessile bacteria constituting the clumps and planktonic bacteria in the supernatant

To identify potential mechanisms of clump formation, we first investigated genes that were differentially expressed in bacterial cells present in clumps associated with host cells and compared them with planktonic cells in the culture medium that were not associated with host cells. Overall RNA expression profiles of bacteria on 5637, PK15, and Caco-2 cell lines clustered based on the subpopulations of adherent bacteria (C1–C3) or planktonic bacteria (S1–S3) and not the host cell line (Fig. 3A). This was consistent with the idea that during bacterial attachment to the host cell, strain 4972 may utilize similar processes of adhesion on different cells. Therefore, RNAseq data of the *E. coli* strain 4972 on the three different cell lines were pooled, and data analysis revealed a total of 622 differentially expressed genes between adherent and planktonic bacteria.

To investigate whether distinct functional classes of DEGs were differentially expressed in clumps compared to planktonic culture, DEGs were assigned to functional categories based on sequence similarity (Fig. 3B). Four functional classes contained significantly more genes that were expressed at a higher level and three classes at a

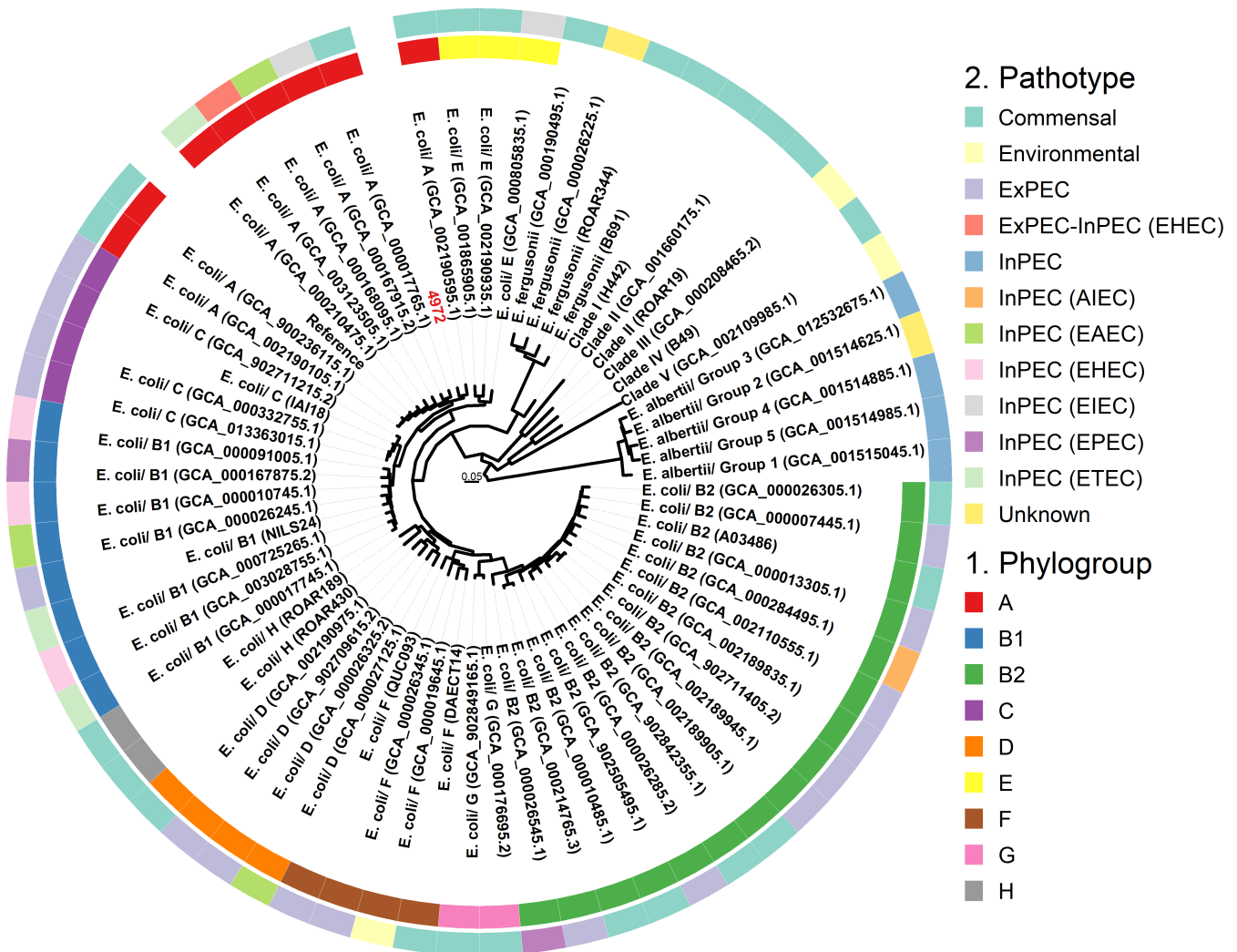


FIG 2 Genomic comparison of *E. coli* strain 4972 phylogroup and pathotypes compared to already published 66 reference genomes of pathogenic and commensal *E. coli* and annotated on the tree with the use of ggtree package. *E. coli* K-12 MG1655 was used as a reference for mapping and snp calling. Placement of the strain 4972 on the tree is highlighted in red.

lower level in clumps compared to planktonic cells. “Molecular chaperones and related functions” class, “Inorganic ion transport and metabolism” class, “Amino acid metabolism and transport” class, and “Coenzyme metabolism” class were enriched in clumps ($P.BH < 0.05$). In contrast, flagellar and fimbriae genes represented by “Secretion, motility, and chemotaxis” COG class and “Nucleotide metabolism and transport” contained significantly more genes that were enriched in bacteria in the supernatants ($P.BH < 0.05$), whereas curli genes (*csgABCDEF*G) remained undifferentiated in the two states. Furthermore, genes that did not fulfill the criteria of any COG category (“NA,” Fig. 3B) were represented by significantly more genes expressed at a lower level in clumps ($P.BH < 0.05$).

Seven genes that exhibited differential expression in clumps compared with planktonic, *ffp*, *pilV*, *spnT*, *yggT*, *spy*, *flgH*, and *fimH*, were chosen for further analysis. The selection was carried on the basis of transcriptomic differential expression, i.e., three genes that were expressed at a higher level and four at a lower level in clumps. Additionally, diverse putative functions and predicted subcellular localization were also taken into account along with other sequence-based analyses (Table 1). In terms of prevalence, the seven targeted genes had variable distribution within pathotypes and commensal *E. coli* (Fig. S2). The *flgH*, *spy*, and *yggT* genes were widely distributed, present in at least 95% of

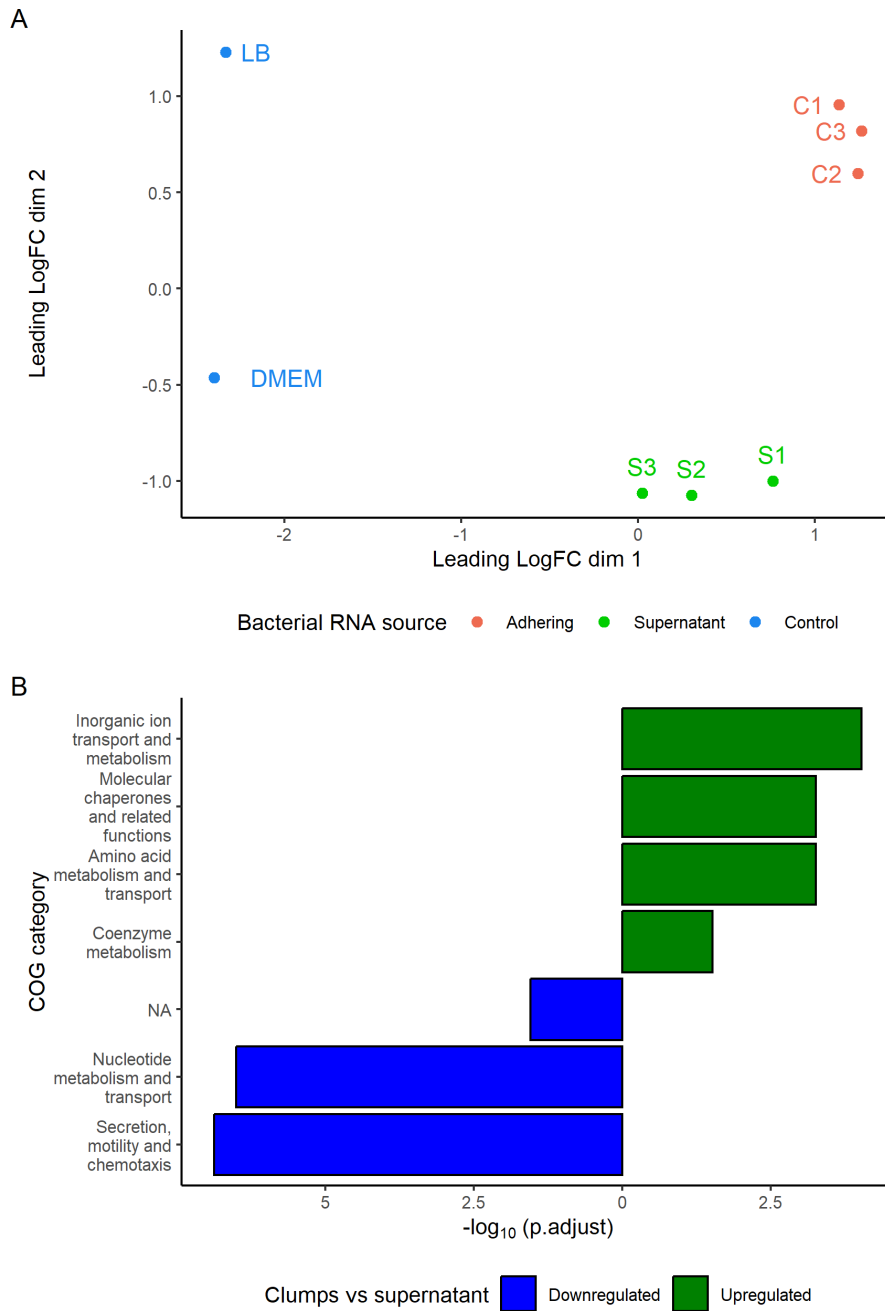


FIG 3 Differential gene expression of adhered and un-adhered bacteria. (A) Multidimensional scaling plot of distances between the samples of RNA sequencing, and leading log-fold change is the root mean square average of the largest log₂-fold changes between each pair of samples. The clumps on the cell lines 5637, PK15, and CaCo-2 are referred to as C1, C2, and C3, respectively. Planktonic bacteria in the supernatant on the cell lines 5637, PK15, and CaCo-2 are labeled as S1, S2, and S3. (B) Differentially expressed significant COGs were compared between adhered and planktonic clumpy *E. coli* for all genes that had *P*.BH value lower than 0.05 in edgeR analysis. Absolute logarithm of Benjamini-Hochberg-adjusted *P* values from GSEA is on the x-axis, and COG categories are shown on the y-axis. NA category corresponds to genes that were not assigned to any of the COGs.

strains in each pathotype. The *ffp* and *fimH* genes were also widely distributed, present in more than 80% of strains in each class, with the exception of DAEC, typical EPEC (tEPEC), and uropathogenic *E. coli* (UPEC) for *ffp*, and EAEC pathotype for *fimH*. The *pilV* and *spnT* genes were rare in all pathotypes of *E. coli*. The *spnT* gene was present in less than 5% of six pathotypes and commensal strains and absent from EHEC, EIEC, and STEC. The *pilV*

TABLE 1 Sequence-based analysis of target genes from transcriptomic data

Gene	Log FC	Putative function	Predicted location	Phobius signal peptide ^a	Phyre2 molecule	TMHMM helices ^a	COG ^b category
<i>ffp</i>	1.59	Fimbrial family protein	Cytoplasm	–	Type-1 fimbrial protein	–	N
<i>pilV</i>	–1.12	Conjugative pilus tip adhesin	Extracellular	n11-22c27/28o	Fimbrial protein	–	N
<i>spnT</i>	–0.88	<i>spnT</i>	Cytoplasm	–	Chaperone of ribosomal protein I4	–	NA
<i>yggT</i>	1.36	Osmotic shock tolerance protein	Cytoplasmic membrane	–	Gamma-aminobutyric acid type b receptor	Yes	S
<i>spy</i>	6.57	<i>spy</i>	Periplasmic	n4-15c23/24o	Periplasmic protein related to spheroblast formation	–	N, P, T
<i>flgH</i>	–2.47	Flagellar basal body L-ring protein	Outer membrane	n14-25c33/34o	Flagellar l-ring protein	–	N
<i>fimH</i>	–2.44	Type 1 fimbrial adhesion	Extracellular and cytoplasmic membrane	n4-14c19/20o	Fimh protein	–	N, U

^a–, not detected.

^bN, secretion, motility, and chemotaxis; O, molecular chaperones and related functions; P, inorganic ion transport and metabolism; T, signal transduction; U, intracellular trafficking, secretion, and vesicular transport; S, unknown function; and NA, unassigned COG category.

gene was present in 17 out of 5,259 (0.32 %) nonpathogenic *E. coli* strains and 1 out of 2,598 (0.038 %) UPEC strains.

Deletion of *flgH* and complementation of *yggT* significantly diminishes adhesion and clump formation on 5637 cells

To investigate the role of various genes on the number of bacteria comprising the clumps when adhered to the 5637 cell line, we constructed isogenic strains in which each gene was individually replaced by a resistance gene. Only $\Delta flgH$ exhibited a significant ($P_{BH} < 0.05$) reduction in adhesion, whereby the mutant's ability to adhere was decreased by up to 80% when compared to the wild type. This mutant was also the only one with visible change, showing reduction in clump formation on the 5637 cell line (Fig. S3). The reintroduction of the *flgH* gene in $\Delta flgH$, i.e., $\Delta flgH$ + pBAD33 *flgH* resulted in significant ($P = 0.05$) regeneration of clumpy adhesion up to 70%, confirming the role of flagella in adhesion. Δffp and its complementation, i.e., Δffp + pBAD33 *ffp* had no significant impact on adhesion. Δspy and $\Delta fimH$ had no significant change in the adhesion ability on the 5637 bladder cell line. Deletion of *yggT* had a statistically significant ($P < 0.05$) increase in adhesion up to 184% that of the wild-type strain (Fig. 4A). $\Delta yggT$ and its respective introduction via pBAD33 system significantly ($P < 0.05$) decreased adhesion to 13%. In the case of $\Delta pilV$, complementation in *trans* did result in a decrease to 53% adhesion, but it was not statistically significant. Similarly, the deletion of *spnT* resulted in an increase to 146% adhesion, but neither the deletion nor complementation of this gene caused a statistically significant change in adhesion to the 5637 cells.

Deletion of *flgH* rendered the strain non-motile, and *ffp* deletion increased motility

Motility is implicated in both biofilm formation and adhesion (36) and to further investigate the role of genes affecting adhesion and the clump phenotype, motility in soft agar was determined (Fig. 4B). Only $\Delta flgH$ lost motility and increased significantly ($P_{BH} < 0.05$) when complemented in *trans*. Δffp had significantly ($P_{BH} < 0.05$) increased motility and when we compared Δffp + pBAD33 with the complemented strain, i.e., Δffp + pBAD33 *ffp*, motility was reduced to the wild-type level ($P_{BH} < 0.05$). Deletion of *pilV*, *spnT*, *yggT*, and *fimH* did not significantly affect motility and remained unchanged throughout the assays (Fig. 4B). The deletion of the *spy* gene caused a significant reduction in motility ($P_{BH} < 0.05$) when compared to the wild-type strain, although the mutant remained motile.

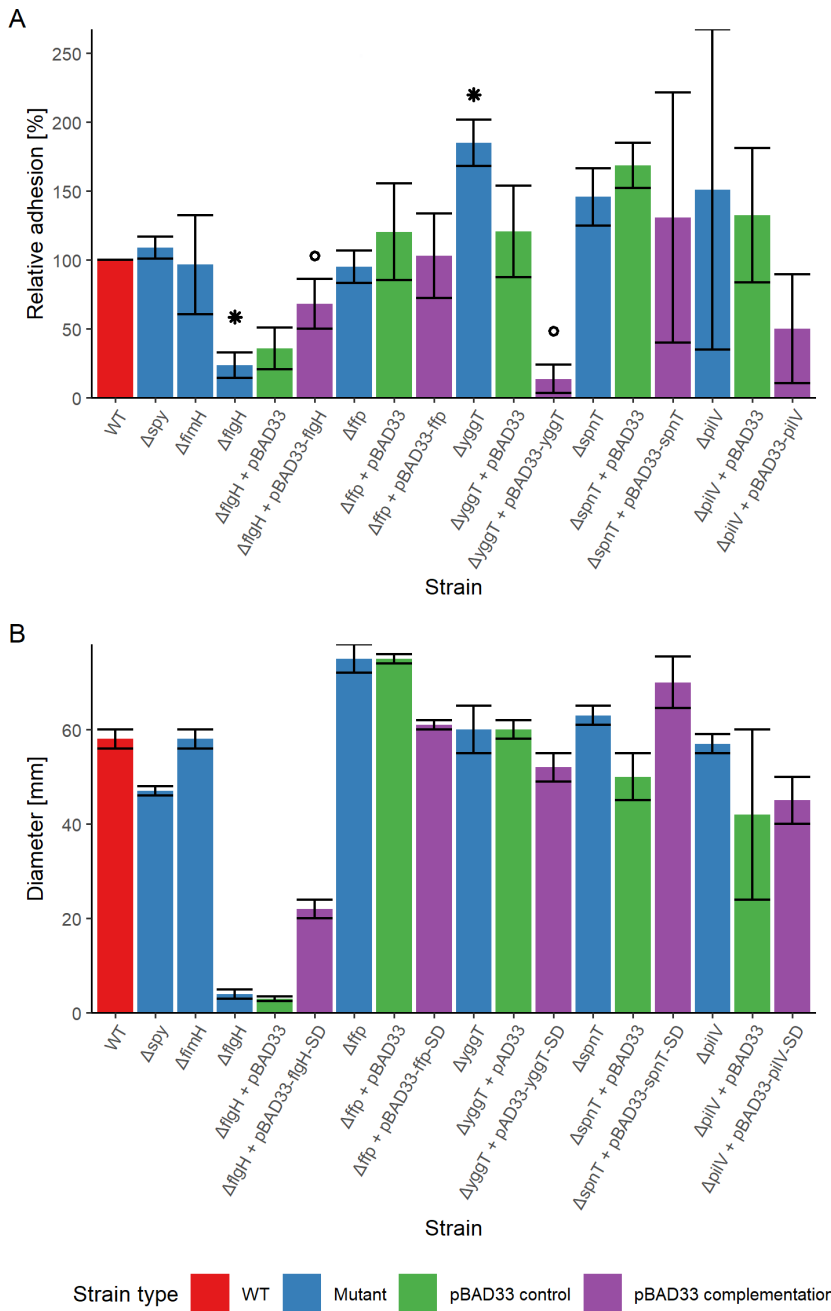


FIG 4 (A) Adhesion and (B) motility assays. (A) Adhesion to 5637 bladder cell line of WT (median value normalized to 100%), gene deletion mutants (Mutant), complemented mutants (pBAD33 complementation), and deletion mutants with empty plasmid (pBAD33 control). The data are shown as median values and median absolute deviation (MAD) of five separate experiments of adhesion assay. X-axis shows the strains, and y-axis depicts relative adhesion (%). Wilcoxon rank-sum test was used to determine the significance between different strains. Statistically significant results indicating changes in mutant adhesion compared to the WT are denoted by an asterisk (*), and significant changes observed upon complementation of the mutant are marked with a small circle (○). (B) Motility of WT (median value normalized to 100%), gene deletion mutants (Mutant), complemented mutants (pBAD33 complementation), and deletion mutants with empty plasmid (pBAD33 control). The data are shown as median values and MAD of three different experiments with three replicates. X-axis shows the strains, and y-axis depicts the diameter of the zone of bacteria. Wilcoxon rank-sum test was used to determine the significance between different strains.

ffp and *spnT* mutants are highly tolerant to ampicillin, and *yggT* mutant is highly susceptible to gentamicin

As the clump formation exhibited similarity to a biofilm-like 3D structure, we addressed the hypothesis that the clump structure protected the bacteria against external stresses. We tested this hypothesis using an adhesion assay in the presence of antibiotic stress. The study not only examined the antibiotic action on the wild-type strain forming the clumpy structure but also elucidated the difference in the survival of viable bacteria of the biofilm-like structure formed by the four gene-deletion mutants, i.e., Δffp , $\Delta yggT$, $\Delta spnT$, and $\Delta pilV$, whose deletion resulted in a general increase in the adhesion percentage. The antibiotics with different modes of action against *E. coli* were selected to stress the clumpy structure. The clumps formed by wild-type strain 4972 were exposed to two different antibiotics, ampicillin and gentamicin, and the concentrations determined were 16,000 and 2 $\mu\text{g}/\text{mL}$, respectively, as depicted by the results of antibiotic dose-ranging experiments (Fig. S4). The exposure to these antibiotic concentrations resulted in a reduction of more than 50% of the viable bacteria in the clumps of the wild-type strain.

In case of ampicillin exposure, clumps produced by Δffp and $\Delta spnT$ were significantly ($P_{\text{BH}} < 0.05$) more resistant and protected the bacterial clumps, with $\Delta spnT$ showing the highest number of viable bacteria in terms of relative adhesion as exhibited in Fig. 5A. It is also surprising to note that the viable bacteria after ampicillin action on the clumpy structures by Δffp and $\Delta spnT$ did not significantly differ from the wild-type strain. It is also important to note that $\Delta spnT$ maintained the best overall bacterial density in media with and without ampicillin, whereas Δffp provided the greatest relative protection from ampicillin (Fig. 5B). In the case of gentamicin stress, clumps built by $\Delta yggT$ were highly susceptible, and no viable bacteria were detected after gentamicin exposure. Both $\Delta spnT$ and $\Delta yggT$ had significantly ($P_{\text{BH}} < 0.05$) low percentage of survival under gentamicin stress when compared to the wild type. In contrast, Δffp and $\Delta pilV$ production of clumpy structure variedly protected the bacteria against the stress, similar to the wild-type strain; therefore, the decrease is not significant (Fig. 5C and D).

Wild-type strain 4972, Δffp , $\Delta yggT$, $\Delta spnT$, and $\Delta pilV$ were also tested for sensitivity against the antibiotic stress in the planktonic state, i.e., in the absence of host eukaryotic cells in MIC assays. Antibiotics inhibited growth at a very low concentration of ampicillin (8 $\mu\text{g}/\text{mL}$) and gentamicin (0.5 $\mu\text{g}/\text{mL}$). Furthermore, the growth declined with an increase in ampicillin concentrations and did not differ between wild type and mutants as shown in Fig. S5A. Stress with gentamicin depicted similar results, where it significantly inhibited the growth as shown in Fig. S5B. Only $\Delta yggT$ was significantly ($P < 0.05$) more susceptible at the gentamicin concentration of 0.25 $\mu\text{g}/\text{mL}$.

Global proteome analysis

In order to add another omics layer, a global proteomics analysis of the wild-type strain was carried out for adhered and planktonic bacteria. The aim was to get a clearer picture of the differential expression of proteins involved in sessile and planktonic populations. As a result of this comparison, 325 proteins were expressed at a higher level in bacteria in the supernatant as compared to 422 proteins having higher expression in the adhering state of bacteria, as shown in the volcano plot (Fig. S6). It was also interesting to note that three genes (i.e., *flgH*, *spy*, and *spnT*) out of seven genes targeted in this study were detected at the proteome level, and two of them (*spy* and *spnT*) were significantly differentially expressed ($P_{\text{BH}} < 0.05$).

COGs were compared between clumps and planktonic bacteria (Fig. 6). The groups of proteins categorized as "Translation, including ribosome structure and biogenesis" were significantly ($P_{\text{BH}} < 0.05$) downregulated group of COG in the adherent bacteria. The proteins categorized as "Carbohydrate metabolism and transport" and "Replication recombination and repair" were significantly enriched ($P_{\text{BH}} < 0.05$). Additionally, specific flagellar proteins such as *flgE* and *fliG* were downregulated in the clumps.

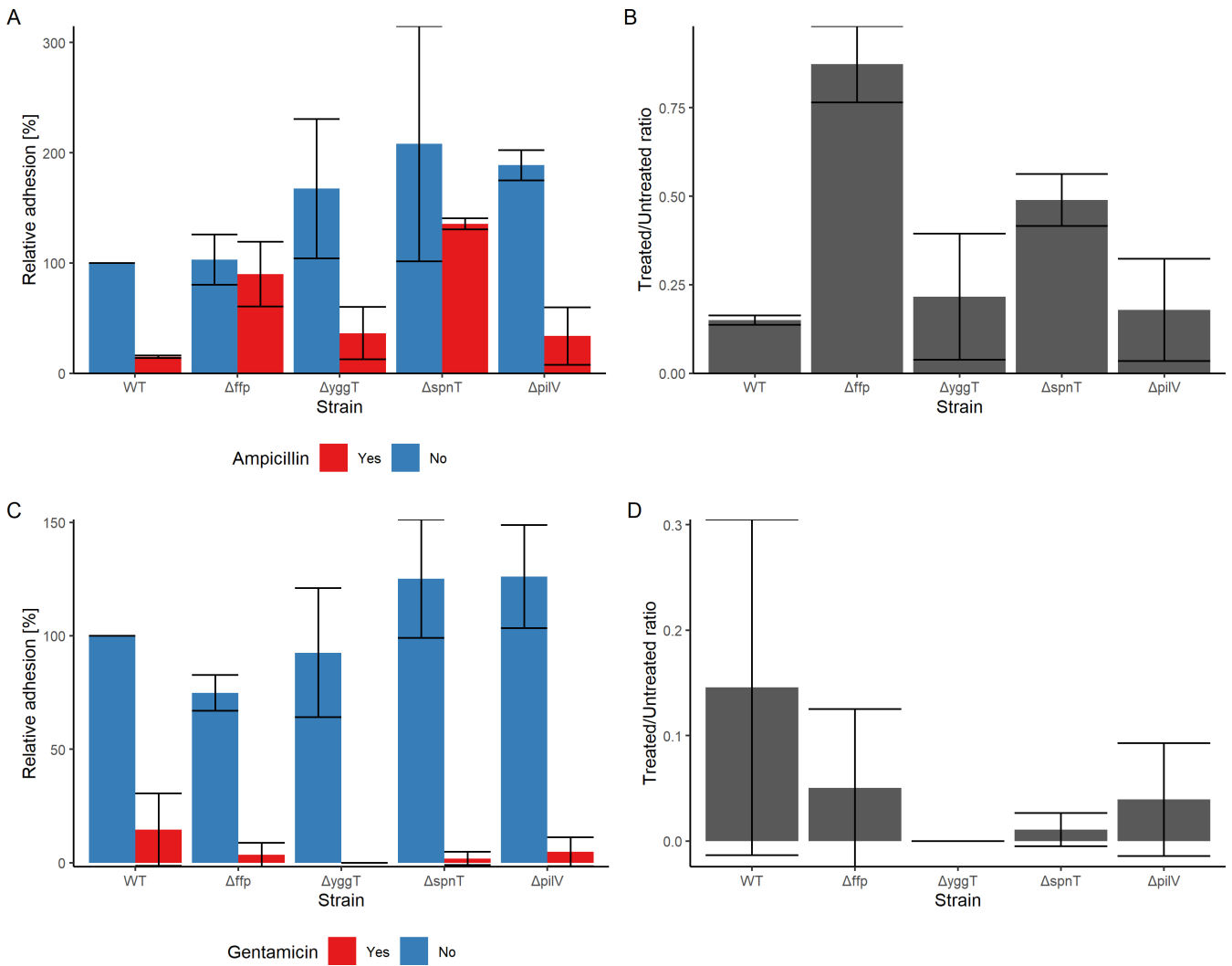


FIG 5 Survival of bacteria in clumps after antibiotic stress. (A) WT without ampicillin stress (median value normalized to 100%), and the four genes are shown with (red bar) and without (blue bar) ampicillin when adhered to the 5637 cell line. (B) Ampicillin-treated/untreated bacteria ratio for WT along with four genes when adhered to the 5637 cell line. (C) WT without gentamicin stress (median value normalized to 100%), and the four genes are shown with (red bar) and without (blue bar) gentamicin when adhered to 5637 cell line. (D) Gentamicin-treated/untreated bacteria ratio for WT along with four genes when adhered to 5637 cell line. The stress concentrations used for ampicillin and gentamicin are 16,000 and 2 μ g/mL, respectively. The data comprise median values and median absolute deviation of five separate experiments in triplicate (A and B) or ratio of median values for antibiotic-treated and untreated bacteria (C and D). The strain names are on the x-axis, and the relative median adhesion is shown on the y-axis. Wilcoxon rank-sum test with $P_{BH} < 0.05$ was used to determine the significance between different strains.

DISCUSSION

The study characterized the biofilm-like 3D structure of clumps formed during adhesion by *E. coli* strain 4972 to different epithelial cells. Normally, biofilm formation is studied on synthetic or plastic materials as it is difficult to quantify them on the living cells or structures. The strain was tested in multiple media to rule out media or synthetic/plastic surfaces, influencing clump formation as the strain strictly formed these structures on the epithelial cells and most prominently on the human urinary bladder cells. We have taken advantage of the unique ability of this strain and characterized it using omics technologies. Clump formation on the cell surface and the effects on the antimicrobial resistance were also studied. Clumps exhibited the characteristics of a biofilm where the adhered or sessile bacterial population on the cell line was shielded from stress by two different antibiotics in contrast to unattached or planktonic bacteria. Hence, the

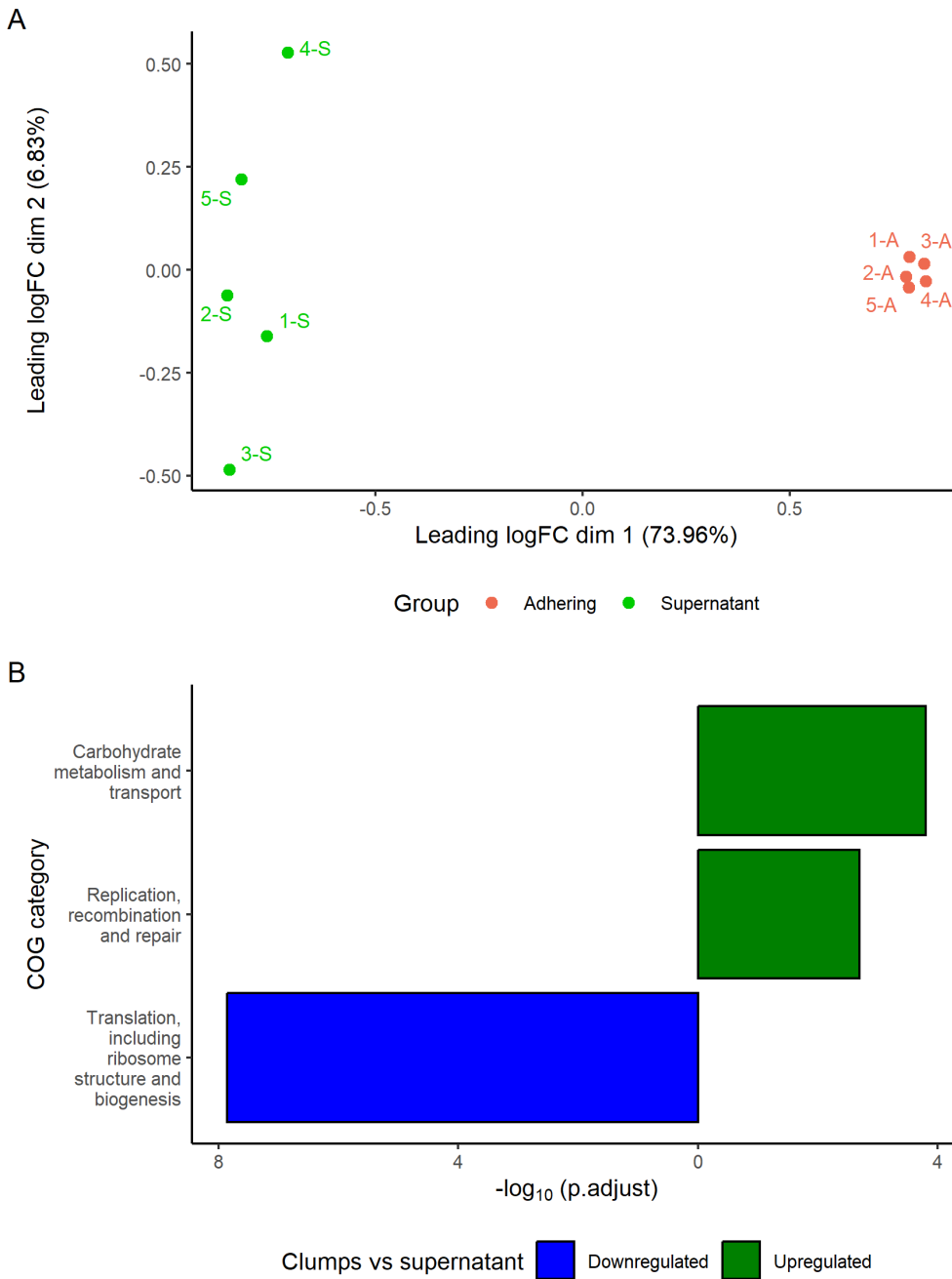


FIG 6 Differential protein expression of adhered clumps and bacteria in the supernatant. (A) Multidimensional scaling plot of distances between the samples, and leading log-fold change is the root mean square average of the largest log₂-fold changes between each pair of samples. (B) Significantly differentially expressed proteins categorized on the genes/proteins in the genome. COGs were compared between adhered and planktonic clumpy *E. coli* for all genes that had *P*.BH value lower than 0.05 and logFC values below -0.5 or above 0.5 in limma analysis. Absolute logarithm of Benjamini-Hochberg adjusted *P* values from GSEA is on the x-axis, and COG categories are shown on the y-axis.

adhesion phenotype, similar to biofilm served as a physical barrier, and the thickness and composition allowed the bacteria to bear a high concentration of antibiotics (42) in contrast to bacteria in the planktonic state, which leaves them vulnerable to antibiotic action. Inherently, the clumpy *E. coli* is sensitive to both antibiotics as shown by the MIC assay. It is important to note that the parameters of the MIC assay and adhesion assay with antibiotic stress are widely different and cannot be compared directly, but it

is necessary to highlight the peculiar behavior of the clumpy adhesion phenotype, which allows the bacteria to survive an extremely high concentration of one antibiotic, i.e., ampicillin (16 mg/mL), while leaving them more susceptible to another, i.e., gentamicin (2 µg/mL) at a lower concentration.

RNAseq data and global proteome analysis comparing unattached and adhered bacteria showed DEGs and differentially abundant proteins in each state. Transcriptomes and proteomes of bacteria in clumps clustered together and diverged from transcriptomes and proteomes of bacteria in the supernatant. Transcriptomes of both pathogenic and nonpathogenic *E. coli* with the ability to form biofilm have been studied earlier (*E. coli* K-12, symptomatic bacteriuria *E. coli*, and ocular *E. coli*) using various platforms. It was reported that overall motility and fimbriae genes are downregulated in the transcriptome when *E. coli* adhere to host cells and form biofilms (43, 44). Both flagella genes (i.e., *flgBCDEFGHIJ*) and fimbriae genes (i.e., *fimACDEFGHI*) were downregulated in the clumps, which is in line with the gene expression of symptomatic bacteriuria *E. coli* strains 83972 and VR50 during biofilm. In both cases, clump formation and biofilm, none of the curli genes, *csgABCDEFG*, were differentially expressed at the transcriptomic level (45). Genes categorized as "Inorganic transport and metabolisms," and "Coenzyme transport and metabolism," were significantly upregulated in clumps, and similar expression pattern was observed in biofilms by *E. coli* K-12 and symptomatic bacteriuria *E. coli*. Post-translational modification, protein turnover, and chaperons were upregulated in clumps, but the expression varied in *E. coli* K-12 but was the opposite in the cases of symptomatic bacteriuria *E. coli* and ocular *E. coli* (45–48).

The study also elucidates the contribution of factors crucial to clump formation to the susceptibility of the selected antibiotics. Based on RNAseq data, gene candidates (i.e., *flgH*, *fimH*, *spy*, *ffp*, *ygjT*, *spnT*, and *pilV*) were selected for gene deletion and characterized. The flagella genes were downregulated at both transcriptomic and proteomic levels. The target gene *flgH* encodes flagellar L-ring protein, which is the basic subunit making up the lipopolysaccharide ring of the flagellar basal body and lies in the outer membrane and encircles the basal body rod of the flagella. L ring is embedded in the cell wall, and the flagellar hook (encoded by *flgE*) does not proceed until the P ring (encoded by *flgI*) and L ring are assembled, rendering the bacteria nonmotile (49). Upon adhesion to the 5637 cell line, $\Delta flgH$ did not develop a clumpy phenotype and had lost the ability to adhere. This is in accordance with other studies where functional flagella were found as the basis for adhesion to host cells. It has also been reported that a strong relationship exists in *E. coli* between motility, flagella, and biofilm architecture (50, 51). Type 1 fimbriae were downregulated at the transcriptomic level and showed the same behavior at the proteome level. However, $\Delta fimH$ exhibited no effect on adhesion and motility. These results may be a consequence of the compensatory effect, in which suppressed expression of one fimbrial cluster may result in the overexpression of other adhesins and consequently maintain the adhesion capacity of the deletion mutant to a cell line or tissue evading change detection (6). Similar to our results, several *E. coli* FimH mutants remained unaltered in adherence phenotype to epithelial cells when compared to parental strains (52, 53).

Protein folding genes were mostly upregulated, with *spy* periplasmic chaperone being significantly upregulated at both transcriptomic and proteomics levels. It encodes an ATP-independent periplasmic chaperone known for preventing protein aggregation and assisting protein refolding and is highly induced in biofilms (46, 54). The evaluation of the comparative assays found that the deletion of *spy* had no effect on clump formation, but it caused a reduction in motility. These results seem to be consistent with previously published research, as it has been shown that periplasmic chaperones seem to be almost functionally redundant; the loss of certain chaperones in deletion mutants could, therefore, be very well compensated, which may explain the lack of change in phenotype (55). In order to fully understand the role of chaperones, more mutants are required to be tested.

The fimbrial family protein or *ffp* gene upregulated in clumps was targeted for deletion. According to the Phyre2 tool, it is a structural protein with similarity to type 1 fimbrial proteins, but no functional information is available for this gene. Interestingly, Δ *ffp* was more motile compared to the wild type, and its overexpression significantly reduced motility. Similar characteristics are exhibited by the deletion of *papX*, another fimbrial gene (56). The overexpression of *papX* correlated with both a significant reduction in flagellin protein synthesis and flagella assembled on the cell surface. *papX* has been implicated in being the master regulator of flagellar synthesis and chemotaxis of UPEC CFT073 (56). When ampicillin stress was applied after the clump formation, Δ *ffp* resisted significantly better compared to the wild type, showing the greatest relative protection from ampicillin. The role of the *ffp* gene requires extensive research in the direction of being a regulator function relating to motility and leading to tolerance of ampicillin's action of interference with cell wall synthesis or permeability in clumps or biofilm-like structures.

Pilus tip adhesin from the shufflon system or *pilV*, a downregulated gene, was identified in the RNAseq analysis. It has been reported that plasmid 1 of the Shiga-toxic hybrid strain 12-05829 carries a type IV pilus biosynthesis locus (*pil*) consisting of 11 genes (*pilL* to *pilV* and *pilI*) (57). Type IV pili are important virulence factors in many pathogenic Gram-negative bacteria, as they mediate adherence to eukaryotic cells and host colonization and are associated with other phenomena related to pathogenicity, including bacterial aggregation, biofilm formation, DNA binding and uptake, and twitching motility (58). Similar to our results of the adhesion assay, deletion of the *pil* locus did not affect bacterial auto-aggregation or adhesion to HEP-2 cells in strain 12-05829 and O113:H2 98NK2 strain (57).

Similar to Δ *ffp*, Δ *spnT* resisted ampicillin stress more than the wild-type strain under the same conditions. A concentration of 16 mg/mL of ampicillin did not exhibit a significant effect on bacteria in the clumps. In contrast, Δ *spnT* was highly susceptible to gentamicin with the lowest number of surviving bacteria in the clumps. Wei et al. (59) failed to disrupt *spnT* but studied its overexpression in *Serratia marcescens* and *E. coli* MG1655. The overexpression of *spnT* in *E. coli* MG1655 resulted in elongated cell morphology and nucleoids after arabinose-mediated induction of *spnT* expression. This indicated that enhanced expression of *spnT* increases the DNA content per *E. coli* cell and blocks the replication cycle in *E. coli*, possibly through the PriA pathway, causing bacterial DNA aggregation and inhibiting cell division. The study by Horng et al. (60) reported that overexpression of this gene in *S. marcescens* adversely affected pigment production and sliding motility. They were also unable to disrupt the gene by using homologous recombination so only reported its overexpression. In our case, it can be hypothesized that since overexpression inhibits cell division, deletion of this gene should have the opposite effect, which was observed by higher numbers of the bacteria in Δ *spnT* clumps. Similarly, overexpression had an insignificant negative effect on adhesion when induced with L-arabinose in the pBAD33 system in agreement with the previously reported study (60). In conclusion, the role of this gene points toward the regulator function, whereby it contributes to the regulation of cell division indirectly affecting adhesion.

Osmotic shock tolerance protein known as YggT is located in the cytoplasmic membrane and possibly forms transmembrane helices (61). The role of YggT in adhesion has not been observed or studied before. It has been reported that *yggT* is one of the genes responsible for phage shock protein (Psp) response in *Yersinia enterocolitica* and *E. coli* (62). *yggT* induced *pspA* expression, but the underlying mechanisms remain elusive. The Psp system, which responds to extracytoplasmic stress, is enhanced by the overexpression of *yggT*. The gene is also associated with signaling upon extracytoplasmic stress. *yggT* has been implicated in antibiotic resistance in *Salmonella enterica*, where *yggT*-deleted D14 mutant showed a decrease in streptomycin resistance (fourfold). Furthermore, complementation and the overexpression of *yggT* in *S. enterica* ATCC 14028 increased the streptomycin resistance (63). In this study, Δ *yggT* in the gentamicin-stressed adhesion assay as well as the gentamicin MIC assay exhibited the highest

sensitivity compared to the wild type and other mutants. This can be explained by the fact that streptomycin and gentamicin are both aminoglycoside antibiotics (64). In the case of ampicillin, $\Delta yggT$ did not significantly alter antibiotic resistance in *S. enterica*, which is in line with our results, where ampicillin stress in adhesion and MIC assay did not significantly vary between wild type and $\Delta yggT$ (63). In conclusion, *yggT*-expressing membrane protein is implicated in adhesion as well as gentamicin susceptibility.

The study opens new areas of research in terms of other genes whose function still needs to be elucidated, as well as other phenotypes of *E. coli* adhesion to various cell lines, organoids, and ultimately in animal models. As it focuses on one aspect, i.e., the adhesion phenotype for antibiotic tolerances, other possibilities have not been taken into account and should also be explored, whereby bacterial physiology, such as reduced permeability and increased efflux (65), host cell invasion (66), and processes associated with the development and maintenance of cellular adhesion and clumping, may also be a contributing factor in sensitivity to the antibiotics. The differences in growth rates in the biofilm-like clumps in comparison to planktonic conditions can also affect the efficiency of antibiotic action, whereby some antibiotics exhibit strong growth rate dependence and slow growth is often associated with decreased susceptibility (67). It has been previously known that bacterial cells embedded in biofilms are extremely difficult to eliminate through antibiotics, as drug penetration is drastically reduced, and the metabolic activity of the bacterial cells is slowed down (68, 69). In terms of future perspectives, first, proteomics data from this study point toward multiple hypothetical and putative proteins whose function and role require further investigation. Second, the study will act as a foundation for testing the hypothesis that various adhesion phenotypes exhibited by pathogenic *E. coli* allow them to survive or tolerate antibiotics and other stresses that offer an advantage and enable effective pathogenesis.

ACKNOWLEDGMENTS

We thank Hermine Kienberger and Franziska Hackbarth for technical assistance and Miriam Abele for mass spectrometric support at BayBioMS.

This work was supported by EPIC-XS (project number 823839) funded by the Horizon 2020 program of the European Union. R.A.K. was supported by the UKRI Institute Strategic Programme Microbes in the Food Chain BB/R012504/1 and its constituent project(s) BBS/E/F/000PR10348 and BBS/E/F/000PR10349.

AUTHOR AFFILIATIONS

¹Institute for Biotechnology, Brandenburg University of Technology (BTU) Cottbus-Senftenberg, Senftenberg, Germany

²Department of Bioinformatics and Genomics, Faculty of Biotechnology, University of Wrocław, Wrocław, Poland

³Department of Biochemistry and Molecular Biology, Faculty of Veterinary Medicine, Wrocław University of Environmental and Life Sciences, Wrocław, Poland

⁴Department of Microbiology, Institute of Genetics and Microbiology, University of Wrocław, Wrocław, Poland

⁵Bavarian Center for Biomolecular Mass Spectrometry (BayBioMS), School of Life Sciences, Technical University Munich (TUM), Munich, Germany

⁶National Institute for Biotechnology and Genetic Engineering College, Pakistan Institute of Engineering and Applied Sciences (NIBGE-C, PIEAS), Faisalabad, Pakistan

⁷Quadram Institute Biosciences, Norwich Research Park, Norwich, United Kingdom

AUTHOR ORCIDs

Muhammad Moman Khan  <http://orcid.org/0000-0003-4493-1563>

Katarzyna Sidorczuk  <http://orcid.org/0000-0001-6576-9054>

Aamir Ali  <http://orcid.org/0000-0002-8433-4328>

Robert A. Kingsley  <http://orcid.org/0000-0002-0194-6485>

Peter Schierack  <http://orcid.org/0000-0002-6445-3506>

Rafał Kolenda  <http://orcid.org/0000-0002-8145-579X>

AUTHOR CONTRIBUTIONS

Muhammad Moman Khan, Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Writing – original draft, Writing – review and editing | Katarzyna Sidorczuk, Investigation, Methodology, Software, Visualization, Writing – review and editing | Juliane Becker, Investigation, Methodology | Adrianna Aleksandrowicz, Methodology, Writing – review and editing | Karolina Baraniewicz, Investigation, Writing – review and editing | Christina Ludwig, Investigation, Methodology, Writing – review and editing | Aamir Ali, Investigation, Writing – review and editing | Robert A. Kingsley, Investigation, Writing – review and editing | Peter Schierack, Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review and editing | Rafał Kolenda, Conceptualization, Formal analysis, Investigation, Methodology, Supervision, Visualization, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

Whole genome sequence for strain 4972 is freely available from the GenBank database under accession number [GCA_018429805.1](https://ncbi.nlm.nih.gov/GenBank/record/GCA_018429805.1) (13). RNA sequences of adhered and unattached bacterial strain 4972 are available in the NCBI BioProject database under accession number [PRJNA868552](https://ncbi.nlm.nih.gov/bioproject/PRJNA868552). The whole genome sequences of all the deletion mutants for this study are available in the NCBI BioProject database under accession number [PRJNA880721](https://ncbi.nlm.nih.gov/bioproject/PRJNA880721). All LC-MS/MS data files and MaxQuant output files have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the data set identifier [PXD042731](https://proteomecentral.proteomexchange.org/protein/PXD042731).

ADDITIONAL FILES

The following material is available [online](#).

Supplemental Material

Supplemental material (Spectrum02606-23-s0001.pdf). Tables S1 to S5; Figures S1 to S6.

REFERENCES

- Bélangier L, Garenaux A, Harel J, Boulianne M, Nadeau E, Dozois CM. 2011. *Escherichia coli* from animal reservoirs as a potential source of human extraintestinal pathogenic *E. coli*. *FEMS Immunol Med Microbiol* 62:1–10. <https://doi.org/10.1111/j.1574-695X.2011.00797.x>
- Poirel L, Madec J-Y, Lupo A, Schink A-K, Kieffer N, Nordmann P, Schwarz S. 2018. Antimicrobial resistance in *Escherichia coli*. *Microbiol Spectr* 6:6. <https://doi.org/10.1128/microbiolspec.ARBA-0026-2017>
- Murray CJL, Ikuta KS, Sharara F, Swetschinski L, Robles Aguilar G, Gray A, Han C, Bisignano C, Rao P, Wool E, et al. 2022. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* 399:629–655. [https://doi.org/10.1016/S0140-6736\(21\)02724-0](https://doi.org/10.1016/S0140-6736(21)02724-0)
- Tacconelli E, Carrara E, Savoldi A, Harbarth S, Mendelson M, Monnet DL, Pulcini C, Kahlmeter G, Kluytmans J, Carmeli Y, Ouellette M, Outtersson K, Patel J, Cavalieri M, Cox EM, Houchens CR, Grayson ML, Hansen P, Singh N, Theuretzbacher U, Magrini N, WHO Pathogens Priority List Working Group. 2018. Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect Dis* 18:318–327. [https://doi.org/10.1016/S1473-3099\(17\)30753-3](https://doi.org/10.1016/S1473-3099(17)30753-3)
- Antão E-M, Ewers C, Gürlebeck D, Preisinger R, Homeier T, Li G, Wieler LH. 2009. Signature-tagged mutagenesis in a chicken infection model leads to the identification of a novel avian pathogenic *Escherichia coli* fimbrial adhesin. *PLoS One* 4:e7796. <https://doi.org/10.1371/journal.pone.0007796>
- Aleksandrowicz A, Khan MM, Sidorczuk K, Noszka M, Kolenda R. 2021. Whatever makes them stick – adhesins of avian pathogenic *Escherichia coli*. *Vet Microbiol* 257:109095. <https://doi.org/10.1016/j.vetmic.2021.109095>
- Govindarajan DK, Viswalingam N, Meganathan Y, Kandaswamy K. 2020. Adherence patterns of *Escherichia coli* in the intestine and its role in pathogenesis. *Med Microecol* 5:100025. <https://doi.org/10.1016/j.medmic.2020.100025>
- Scaletsky IC, Silva ML, Trabulsi LR. 1984. Distinctive patterns of adherence of enteropathogenic *Escherichia coli* to HeLa cells. *Infect Immun* 45:534–536. <https://doi.org/10.1128/iai.45.2.534-536.1984>
- Ballén V, Cepas V, Ratia C, Gabasa Y, Soto SM. 2022. Clinical *Escherichia coli*: from biofilm formation to new antibiofilm strategies. *Microorganisms* 10:1103. <https://doi.org/10.3390/microorganisms10061103>
- Frömmel U, Böhm A, Nitschke J, Weinreich J, Groß J, Rödiger S, Wex T, Ansoerge H, Zinke O, Schröder C, Roggenbuck D, Schierack P. 2013. Adhesion patterns of commensal and pathogenic *Escherichia coli* from humans and wild animals on human and porcine epithelial cell lines. *Gut Pathog* 5:31–31. <https://doi.org/10.1186/1757-4749-5-31>

11. Seemann T. 2017. Shovill: Faster (SPAdes) assembly of (Illumina) reads. <https://github.com/tseemann/shovill>.
12. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>
13. Kolenda R, Sidorczuk K, Noszka M, Aleksandrowicz A, Khan MM, Burdukiewicz M, Derek P, Schierack P. 2021. Genome placement of alpha-hemolysin defines virulence-associated factor profile of *Escherichia coli*. *Microb Genom* 7:000743. <https://doi.org/10.1099/mgen.0.000743>
14. Denamur E, Clermont O, Bonacorsi S, Gordon D. 2021. The population genetics of pathogenic *Escherichia coli*. *Nat Rev Microbiol* 19:37–54. <https://doi.org/10.1038/s41579-020-0416-x>
15. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>
16. Yu G. 2020. Using ggtree to visualize data on tree-like structures. *Curr Protoc Bioinformatics* 69:e96. <https://doi.org/10.1002/cpbi.96>
17. Schiebel J, Böhm A, Nitschke J, Burdukiewicz M, Weinreich J, Ali A, Roggenbuck D, Rödiger S, Schierack P. 2017. Genotypic and phenotypic characteristics associated with biofilm formation by human clinical *Escherichia coli* isolates of different pathotypes. *Appl Environ Microbiol* 83:e01660-17. <https://doi.org/10.1128/AEM.01660-17>
18. Stead MB, Agrawal A, Bowden KE, Nasir R, Mohanty BK, Meagher RB, Kushner SR. 2012. RNAsnap: a rapid, quantitative and inexpensive, method for isolating total RNA from bacteria. *Nucleic Acids Res* 40:e156. <https://doi.org/10.1093/nar/gks680>
19. Wingett SW, Andrews S. 2018. FastQ Screen: a tool for multi-genome mapping and quality control. *F1000Res* 7:1338. <https://doi.org/10.12688/f1000research.15931.2>
20. Liao Y, Smyth GK, Shi W. 2019. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res* 47:e47. <https://doi.org/10.1093/nar/gkz114>
21. Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–140. <https://doi.org/10.1093/bioinformatics/btp616>
22. McCarthy DJ, Chen Y, Smyth GK. 2012. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res* 40:4288–4297. <https://doi.org/10.1093/nar/gks042>
23. Wickham H, Chang W, Henry L, Pedersen TL, Takahashi K, Wilke C, Woo K, Yutani H, RStudio. 2019. ggplot2: create elegant data visualisations using the grammar of graphics (3.2.0). <https://CRAN.R-project.org/package=ggplot2>.
24. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 10:845–858. <https://doi.org/10.1038/nprot.2015.053>
25. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, Dao P, Sahinalp SC, Ester M, Foster LJ, Brinkman FSL. 2010. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* 26:1608–1615. <https://doi.org/10.1093/bioinformatics/btq249>
26. Käll L, Krogh A, Sonnhammer ELL. 2007. Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. *Nucleic Acids Res* 35:W429–W432. <https://doi.org/10.1093/nar/gkm256>
27. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J Mol Biol* 305:567–580. <https://doi.org/10.1006/jmbi.2000.4315>
28. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, von Mering C, Bork P. 2019. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* 47:D309–D314. <https://doi.org/10.1093/nar/gky1085>
29. Galperin MY, Makarova KS, Wolf YI, Koonin EV. 2015. Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res* 43:D261–D269. <https://doi.org/10.1093/nar/gku1223>
30. Tatusov RL, Galperin MY, Natale DA, Koonin EV. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res* 28:33–36. <https://doi.org/10.1093/nar/28.1.33>
31. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, Fu X, Liu S, Bo X, Yu G. 2021. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation (Camb)* 2:100141. <https://doi.org/10.1016/j.xinn.2021.100141>
32. Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 97:6640–6645. <https://doi.org/10.1073/pnas.120163297>
33. Deros V, Deboeck F, Hernalsteens J-P, De Greve H. 2011. Reproducible gene targeting in recalcitrant *Escherichia coli* isolates. *BMC Res Notes* 4:213–213. <https://doi.org/10.1186/1756-0500-4-213>
34. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
35. Grzymajlo K, Ugorski M, Suchanski J, Kedzierska AE, Kolenda R, Jarzab A, Biernatowska A, Schierack P. 2017. The novel type 1 fimbriae FimH receptor calreticulin plays a role in *Salmonella* host specificity. *Front Cell Infect Microbiol* 7:326. <https://doi.org/10.3389/fcimb.2017.00326>
36. Ali A, Kolenda R, Khan MM, Weinreich J, Li G, Wieler LH, Tedin K, Roggenbuck D, Schierack P. 2020. Novel avian pathogenic *Escherichia coli* genes responsible for adhesion to chicken and human cell lines. *Appl Environ Microbiol* 86:e01068-20. <https://doi.org/10.1128/AEM.01068-20>
37. Sidorczuk K, Aleksandrowicz A, Burdukiewicz M, Kingsley RA, Kolenda R. 2023. Genomic characterization of enterohaemolysin-encoding haemolytic *Escherichia coli* of animal and human origin. *Microb Genom* 9:mgen000999. <https://doi.org/10.1099/mgen.0.000999>
38. Tyanova S, Temu T, Cox J. 2016. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* 11:2301–2319. <https://doi.org/10.1038/nprot.2016.136>
39. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. 2014. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* 13:2513–2526. <https://doi.org/10.1074/mcp.M113.031591>
40. Team RC. 2020. R: a language and environment for statistical computing. R foundation for statistical computation, Vienna, Austria.
41. Wickham H. 2016. ggplot2: elegant graphics for data analysis. Springer.
42. Yin W, Wang Y, Liu L, He J. 2019. Biofilms: the microbial “Protective Clothing” in extreme environments. *Int J Mol Sci* 20:3423. <https://doi.org/10.3390/ijms20143423>
43. Bayramoglu B, Toubiana D, Gillor O. 2017. Genome-wide transcription profiling of aerobic and anaerobic *Escherichia coli* biofilm and planktonic cultures. *FEMS Microbiol Lett* 364:6. <https://doi.org/10.1093/fems/lnx006>
44. Pratt LA, Kolter R. 1998. Genetic analysis of *Escherichia coli* biofilm formation: roles of flagella, motility, chemotaxis and type I pili. *Mol Microbiol* 30:285–293. <https://doi.org/10.1046/j.1365-2958.1998.01061.x>
45. Hancock V, Klemm P. 2007. Global gene expression profiling of asymptomatic bacteriuria *Escherichia coli* during biofilm growth in human urine. *Infect Immun* 75:966–976. <https://doi.org/10.1128/IAI.01748-06>
46. Beloin C, Valle J, Latour-Lambert P, Faure P, Kzreminski M, Balestrino D, Haagensen JAJ, Molin S, Prensier G, Arbeille B, Ghigo J-M. 2004. Global impact of mature biofilm lifestyle on *Escherichia coli* K-12 gene expression. *Mol Microbiol* 51:659–674. <https://doi.org/10.1046/j.1365-2958.2003.03865.x>
47. Domka J, Lee J, Bansal T, Wood TK. 2007. Temporal gene-expression in *Escherichia coli* K-12 biofilms. *Environ Microbiol* 9:332–346. <https://doi.org/10.1111/j.1462-2920.2006.01143.x>
48. Ranjith K, Arunasri K, Reddy GS, Adicherla H, Sharma S, Shivaji S. 2017. Global gene expression in *Escherichia coli*, isolated from the diseased ocular surface of the human eye with a potential to form biofilm. *Gut Pathog* 9:15. <https://doi.org/10.1186/s13099-017-0164-2>
49. Berg HC. 2003. The rotary motor of bacterial flagella. *Annu Rev Biochem* 72:19–54. <https://doi.org/10.1146/annurev.biochem.72.121801.161737>
50. Kleta S, Nordhoff M, Tedin K, Wieler LH, Kolenda R, Oswald S, Oelschlaeger TA, Bleiss W, Schierack P. 2014. Role of F1C fimbriae, flagella, and secreted bacterial components in the inhibitory effect of probiotic

- Escherichia coli* Nissle 1917 on atypical enteropathogenic *E. coli* infection. *Infect Immun* 82:1801–1812. <https://doi.org/10.1128/IAI.01431-13>
51. Wood TK, González Barrios AF, Herzberg M, Lee J. 2006. Motility influences biofilm architecture in *Escherichia coli*. *Appl Microbiol Biotechnol* 72:361–367. <https://doi.org/10.1007/s00253-005-0263-8>
 52. Hull RA, Donovan WH, Del Terzo M, Stewart C, Rogers M, Darouiche RO. 2002. Role of type 1 fimbria- and P fimbria-specific adherence in colonization of the neurogenic human bladder by *Escherichia coli*. *Infect Immun* 70:6481–6484. <https://doi.org/10.1128/IAI.70.11.6481-6484.2002>
 53. Katani R, Kudva IT, Srinivasan S, Stasko JB, Schilling M, Li L, Cote R, DebRoy C, Arthur TM, Sokurenko EV, Kapur V. 2021. Strain and host-cell dependent role of type-1 fimbriae in the adherence phenotype of super-shed *Escherichia coli* O157:H7. *Int J Med Microbiol* 311:151511. <https://doi.org/10.1016/j.ijmm.2021.151511>
 54. Koldewey P, Stull F, Horowitz S, Martin R, Bardwell JCA. 2016. Forces driving chaperone action. *Cell* 166:369–379. <https://doi.org/10.1016/j.cell.2016.05.054>
 55. Sklar JG, Wu T, Kahne D, Silhavy TJ. 2007. Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in *Escherichia coli*. *Genes Dev* 21:2473–2484. <https://doi.org/10.1101/gad.1581007>
 56. Simms AN, Mobley HLT. 2008. PapX, a P fimbrial operon-encoded inhibitor of motility in uropathogenic *Escherichia coli*. *Infect Immun* 76:4833–4841. <https://doi.org/10.1128/IAI.00630-08>
 57. Lang C, Fruth A, Holland G, Laue M, Mühlen S, Dersch P, Flieger A. 2018. Novel type of pilus associated with a shiga-toxicogenic *E. coli* hybrid pathovar conveys aggregative adherence and bacterial virulence. *Emerg Microbes Infect* 7:203. <https://doi.org/10.1038/s41426-018-0209-8>
 58. Xicohtencatl-Cortes J, Monteiro-Neto V, Ledesma MA, Jordan DM, Francetic O, Kaper JB, Puente JL, Girón JA. 2007. Intestinal adherence associated with type IV pili of enterohemorrhagic *Escherichia coli* O157:H7. American Society for Clinical Investigation. Available from: <https://www.jci.org/articles/view/30727/cite>. Retrieved 06 Jan 2023.
 59. Wei J-R, Soo P-C, Horng Y-T, Hsieh S-C, Tsai Y-H, Swift S, Withers H, Williams P, Lai H-C. 2006. Regulatory roles of *spnT*, a novel gene located within transposon TnTR. *Biochem Biophys Res Commun* 348:1038–1046. <https://doi.org/10.1016/j.bbrc.2006.07.152>
 60. Horng Y-T, Deng S-C, Daykin M, Soo P-C, Wei J-R, Luh K-T, Ho S-W, Swift S, Lai H-C, Williams P. 2002. The LuxR family protein SpnR functions as a negative regulator of *N*-acylhomoserine lactone-dependent quorum sensing in *Serratia marcescens*. *Mol Microbiol* 45:1655–1671. <https://doi.org/10.1046/j.1365-2958.2002.03117.x>
 61. Ito T, Uozumi N, Nakamura T, Takayama S, Matsuda N, Aiba H, Hemmi H, Yoshimura T. 2009. The implication of YggT of *Escherichia coli* in osmotic regulation. *Biosci Biotechnol Biochem* 73:2698–2704. <https://doi.org/10.1271/bbb.90558>
 62. Maxson ME, Darwin AJ. 2004. Identification of inducers of the *Yersinia enterocolitica* phage shock protein system and comparison to the regulation of the RpoE and Cpx extracytoplasmic stress responses. *J Bacteriol* 186:4199–4208. <https://doi.org/10.1128/JB.186.13.4199-4208.2004>
 63. Liang Y, Deng F, Mu P, Wen J, Deng Y. 2021. Quantitative proteomics implicates YggT in streptomycin resistance in *Salmonella enterica* serovar Enteritidis. *Biotechnol Lett* 43:919–932. <https://doi.org/10.1007/s10529-021-03083-4>
 64. Bryan LE, Van Den Elzen HM. 1977. Effects of membrane-energy mutations and cations on streptomycin and gentamicin accumulation by bacteria: a model for entry of streptomycin and gentamicin in susceptible and resistant bacteria. *Antimicrob Agents Chemother* 12:163–177. <https://doi.org/10.1128/AAC.12.2.163>
 65. Darby EM, Trampari E, Siasat P, Gaya MS, Alav I, Webber MA, Blair JMA. 2023. Molecular mechanisms of antibiotic resistance revisited. *Nat Rev Microbiol* 21:280–295. <https://doi.org/10.1038/s41579-022-00820-y>
 66. Lewis AJ, Richards AC, Mulvey MA. 2016. Invasion of host cells and tissues by uropathogenic bacteria. *Microbiol Spectr* 4. <https://doi.org/10.1128/microbiolspec.UTI-0026-2016>
 67. Greulich P, Scott M, Evans MR, Allen RJ. 2015. Growth-dependent bacterial susceptibility to ribosome-targeting antibiotics. *Mol Syst Biol* 11:796. <https://doi.org/10.15252/msb.20145949>
 68. Stokes JM, Lopatkin AJ, Lobritz MA, Collins JJ. 2019. Bacterial metabolism and antibiotic efficacy. *Cell Metab* 30:251–259. <https://doi.org/10.1016/j.cmet.2019.06.009>
 69. Ito A, Taniuchi A, May T, Kawata K, Okabe S. 2009. Increased antibiotic resistance of *Escherichia coli* in mature biofilms. *Appl Environ Microbiol* 75:4093–4100. <https://doi.org/10.1128/AEM.02949-08>