Phylogenetic background and extraintestinal virulence genotypes of *Escherichia coli* vaginal strains isolated from adult women

Genotipurile de virulentă extraintestinală și încadrarea filogenetică a tulpinilor vaginale de *Escherichia coli* izolate de la femei adulte

Codruța-Romanita Usein1*, Laura Andrea Grigore2, Raluca Margareta Georgescu2, Mădălina Cornelia Băltoiu1, Maria Condei1, Monica Delia Teleman3

1. Laboratory of Molecular Epidemiology, National Institute of Research-Development for Microbiology and Immunology “Cantacuzino”, Bucharest, Romania
2. Laboratory of Medical Analyses, National Institute of Research-Development for Microbiology and Immunology “Cantacuzino”, Bucharest, Romania
3. Department of Microbiology and Epidemiology, University of Medicine and Pharmacy “Carol Davila”, Bucharest, Romania

Abstract

Distinctive *Escherichia coli* clones, designated as extraintestinal pathogenic *E. coli* (ExPEC), are responsible for extraintestinal infections. Human vagina may be a transient residence for *E. coli* and its immediate source for adjoining sites such as the urinary tract. The aim of this study was to complement the routine microbiological diagnostic by examining whether the *E. coli* isolates recovered from vaginal specimens possessed ExPEC-associated traits. Ninety-three *E. coli* isolates were investigated using PCR-based protocols for their phylogenetic origin and virulence genotype, targeting 14 virulence genes. Genetic relatedness among the isolates was assessed using PFGE of XbaI macrorestriction DNA fragments. Phylogenetic groups A, B1, B2, and D accounted for 20%, 2%, 65%, and 13%, respectively. At least one virulence gene region was detected in each examined isolate, their prevalence ranging from 0% (papGI) to 98% (fimH). Group B2-derived isolates, found as diverse based on the PFGE profiles, exhibited the highest virulence content. Overall, certain of the sought genes, e.g. fyuA (85%), irp2 (85%), sfa/focDE (53%), hly (52%), iucC (51%), cnf1 (46%), and papC (42%) were more frequently detected in the examined isolates than others, e.g. sat (17%), ibeA (17%), and afaC (5%). As for papG alleles, papGIII occurred more frequently than papGII. The genetic analysis allowed to distinguish among the vaginal *E. coli* isolates those which possessed an assortment of virulence genes that can promote an infectious process. However, much further work remains to be done before deciding whether this can become feasible in laboratory routine practice.

Keywords: vaginal *E. coli*, ExPEC, extraintestinal virulence genes.

*Corresponding author: Dr. Codruta-Romanita Usein, National Institute of Research-Development for Microbiology and Immunology “Cantacuzino”, Laboratory of Molecular Epidemiology, Splaiul Independentei 103, Sector 5, 050096, Bucharest, Romania
Tel: +40215287223, Fax: +40213184414, E-mail: rusein@cantacuzino.ro
Rezumat

Infectiile extraintestinale cu Escherichia coli se datorează unor clone distincte, reunite sub denumirea de E. coli cu patogenitate extraintestinală (ExPEC). Vaginul poate adăposti temporar E. coli, devenind astfel sursa de contaminare pentru regiuni învecinate, precum tractul urinar. Acest studiu și-a propus să completeze diagnosticul microbiologic curent, urmărind să evidențieze caracteristicile de patogenitate extraintestinală la tulpiii de E. coli izolate din vagin. Nouăzeci și trei de izolate au fost examineate, urmărindu-se originea lor filogenetică și genotipurile de virulență, reprezentat prin 14 gene, folosind protocoloale bazate pe tehnica PCR. Gradul de înrudire genetică s-a apreciat pe baza profilului de fragmente de macrorestricții genomică cu enzima XbaI, obținut prin electroforeza în câmp pulsator (PFGE). Distribuția izolatelor de E. coli în grupuri filogenetice a fost urmată: 20% s-au înregistrat în grupul A, 2% în grupul B1, 65% în grupul B2 și 13% în grupul D. În fiecare dintre izolate s-a detectat cel puțin una dintre regiunile codante pentru virulență, prevalența lor fiind cuprinse între 0% (papGII) și 98% (fimH). Grupul B2, reușind izolate cu un profil PFGE variat, a concentrat cele mai multe gene de virulență. Unele gene de virulență au fost detectate mai frecvent (85% fyuA, 85% irp2, 53% sfa/focDE, 52% hly, 51% iucC, 46% cnf1, 42% papC) decât altele (17% sat, 17% ibeA, 5% afaC). Alela papGIII a fost evidențiată mai frecvent decât papGII. Analiza genetică a permis identificarea izolatelor de E. coli dotează cu potențialul de virulență necesară declanșării unui proces infectios. Pentru ca această investigație să intre în rutina practică de laborator sunt necesare studii suplimentare.

Cuvinte cheie: E. coli vaginal, ExPEC, gene de virulență extraintestinală

Introduction

Extraintestinal pathogenic E. coli (ExPEC) represents a distinct group of pathogenic E. coli that causes most of the extraintestinal E. coli infections (1). ExPEC strains are genetically distinct from commensal E. coli found in the intestinal flora. They are usually characterized by a predominance of phylogenetic group B2 and diverse virulence factors, among which adhesins, toxins, iron sequestration systems, and polysaccharide coatings. By virtue of these specialized virulence factors, ExPEC are able to colonize key anatomical sites outside the host intestinal tract and cause disease (2).

In healthy women, the “bad” extraintestinal E. coli strains are frequently associated with urinary tract infections (UTIs) (2). Most of the UTIs develop in an ascending manner, with E. coli gaining access to the bladder via the urethra, and the initial colonization of the vaginal mucosa is considered a critical step toward infection (3). Sexual transmission of E. coli was documented, the intercourse facilitating not only the introduction of the pathogen to the bladder but also its transfer among partners (4). In addition, in pregnant women, the ExPEC from the vaginal flora may be transmitted to the newborns during or before the delivery leading to neonatal infections, such as meningitis and sepsis (5).

The development of an infection is a complex process which depends on both the infectious agent’s virulence traits and the host’s susceptibility determinants. The transitory presence of E. coli in the vaginal flora tends to be overlooked in the absence of overt symptoms of infection. However, a more comprehensive picture of these vaginal isolates is necessary in order to assess their potential to break the barriers between commensalism and pathogenicity and to consider adopting therapeutic or preventive medical strategies in a particular clinical circumstance. Therefore, this study aimed to reveal ExPEC-associated characteristics in E. coli isolates obtained from the vaginal specimens of adult women before developing overt symptoms. The more refined analysis of the phylogenetic origin and virulence genotype of these isolates served to complement the routine diagnostic based on detection and identification at species level which is usually insufficient for such a heterogeneous species.
**Material and methods**

*E. coli strains*

A total of 93 *E. coli* strains were analyzed. The strains were recovered from the vaginal swab samples of adult females yielding *E. coli* as either sole or dominating microorganism isolated. From the *E. coli* growth, only one colony (isolate) per specimen was selected for molecular investigations. The vaginal samples were collected between May to August 2009, in the Laboratory of Medical Analyses, "Cantacuzino" National Institute of Research-Development for Microbiology and Immunology, from subjects aged between 18 to 74 years (mean age of 33.8 years) with no symptoms of urinary tract infection at the time of the microbiological investigation.

**Phylogenetic classification and virulence genotyping**

All strains were assigned to one of the four major *E. coli* phylogenetic groups (A, B1, B2, and D) as described by Clermont et al. (6), based on the presence of two genes (*chuA* and *yjaA*) and a DNA fragment (TSPE4.C2). The vaginal isolates were screened by PCR for the following virulence-associated genes: *fimH*, *papC*, *sfa/focDE*, *afaC*, *hly*, *cnf1*, *sat*, *iucC*, *fyuA*, *irp2*, and *ibeA*. Strains that were positive for *papC* were tested for alleles of *papG*: *papGI*, *papGII*, and *papGIII*. The primers’ sequences have been previously published (7-9). The gene regions sought were considered as predictors for the following virulence factors: type 1 fimbriae (*fimH*), P fimbriae (*papC* and *papG* alleles), S and/or F1C fimbriae (*sfa/focDE*), Dr family of adhesins (*afaC*), alpha-hemolysin (*hly*), cytotoxic necrotizing factor 1 (*cnf1*), secreted autotransporter toxin (*sat*), aerobactin (*iucC*), invasion of brain endothelium factor (*ibeA*), and *Yersinia* High-Pathogenicity Island (*fyuA*, *irp2*).

A virulence score was calculated as the sum of all virulence genes for which an isolate tested positive, with *papC* and *papG* alleles counting collectively as a single trait. The operational definition of ExPEC isolates in the present study was based on the criterion of "virulence score ≥ 4".

**Pulsed-field gel electrophoresis**

Pulsed-field gel electrophoresis (PFGE) of DNA macrorestriction fragments was used to assess the genetic relatedness of selected strains. The protocol was in concordance with the standardized PulseNet protocol for subtyping *E. coli* O157:H7 by PFGE (10). XbaI-digested DNA fragments were resolved in a CHEF Mapper apparatus (Bio-Rad, Hercules, CA) and the TIFF images of PFGE patterns were analyzed with Fingerprinting II software (BioRad). Similarities of fragments between strains were compared by using Dice coefficient at 1.0% tolerance and 0.5% optimization. A dendrogram was constructed with the unweighted pair-group method with arithmetic averages (UPGMA) clustering method. The isolates were considered clonal if they exhibited indistinguishable profiles (Dice similarity coefficient 100%).

**Statistical analysis**

Proportions were compared using the chi-square test. Virulence scores were compared using the non-parametric Wilcoxon Rank–Sum Test (Mann-Whitney U test). The threshold for statistical significance was P < 0.05. The statistical analysis was performed by using SPSS 11.0 software.

**Results**

**Prevalence of virulence markers and virulence genotypes**

Of the gene regions sought, all but *papG* allele I were detected, ranging in prevalence from 5% (*afaC*) to 98% (*fimH*) (Table 1). The most prevalent virulence genes, in descending order, were: *fimH* (98%), *fyuA* and *irp2* (85% each), followed by *sfa/focDE* (53%), *hly* (52%), and *iucC* (51%) genes. The genes *cnf1* and *papC* occurred in 48% and 43% of the isolates, respectively. All the *cnf1*-positive isolates also harbored *hly* gene. Of the *papC*-positive isolates, 23 isolates contained *papGIII* and 18 isolates *papGII* allele. The combination of *papG* II+III al-
leles was detected in 2 isolates. The genes *ibeA* and *sat* occurred with the same frequency (17%).

Overall, the PCR-based virulence genotyping indicated that the vaginal *E. coli* strains carried between one to ten of the eleven virulence-associated genes selected as targets in this study. There were only 5 isolates harboring one virulence gene, 7 isolates harboring 2 virulence genes, 6 isolates harboring 3 genes, and the rest harbored at least four of the genes. A total of 35 virulence genotypes were identified, of which the following gene associations occurred more frequently: *papC/GIII+sfa/focDE+hly+cnf1+fyuA+irp2+fimH* (12 isolates), *sfa/focDE+hly+cnf1+fyuA+irp2+fimH* (10 isolates), *iucC* (*fyuA+irp2+fimH* (7 isolates), *papC/GIII+sfa/focDE+hly+cnf1+ibeA+fyuA+irp2+fimH* (7 isolates).

**Phylogenetic structure**

The four major *E. coli* phylogenetic groups differed in prevalence, with group B2 accounting for 65% of the studied population versus 20% for group A, 13% for group D, and 2% for group B1.

**Phylogenetic distribution of virulence markers**

The virulence markers selected for the characterization of the vaginal *E. coli* isolates were found to exhibit distinctive and diverse
patterns of phylogenetic distribution. Confine-
ment to group B2 was observed for sfa/focDE
and cnf1 (for both the p values < 10^{-8}). Occur-
rence of virulence markers in two phylogenetic
Groups and significant association with one of
them was seen for hly, ibeA and sat, of which
the first two were more prevalent within group
B2 and the latter in group D. Other virulence
markers were more dispersed (e.g. papC, fimH,
afaC, iucC, fyuA, irp2), occurring in ≥3 phylo-
genetic groups with different prevalences. Al-
though some of them were still significantly as-
sociated with phylogenetic group B2 (e.g.
papC, fyuA and irp2) or group D (e.g. iucC),
others had a broad and homogeneous distribu-
tion in all four major groups (e.g. fimH).

Most individual virulence markers were
concentrated within group B2. We found that
83% isolates from phylogenetic group B2 have
≥ 6 virulence markers as compared with 3% from
Group A, B1 and D taken together (OR = 160; 95% CI 19.06 - 633.61; p < 0.001).
Accordingly, the score derived was tested by
Wilcoxon Rank-Sum Test, demonstrating that
the median aggregate virulence score was
highest for the group B2, whilst the other
groups had significantly lower and comparable
virulence scores (Table 2).

**PFGE typing**

The group B2-derived *E. coli* isolates
were selected to be subjected to PFGE using
XbaI-restricted total DNA. With the exception
of one isolate that consistently sheared during
the PFGE analysis, the isolates yielded inter-
pretable PFGE profiles with XbaI, which were
compared by creating an UPGMA-based simi-
arity dendrogram. This dendrogram revealed 59
distinctive profiles (Figure 1).

### Discussion

The niche of commensal *E. coli* is the
mucous layer of the colon. However, there are
*E. coli* clones that are distinct from the intesti-
nal commensal *E. coli*, possessing specific fit-
tness and virulence attributes which allow adap-
tation to other niches (e.g. urinary tract, central
nervous system, blood) and confer enhanced
ability to cause a broad spectrum of disease in
extraintestinal sites. These virulent clones make
up the ExPEC population of which the uropath-
ogenic *E. coli* (UPEC) group is the most com-
monly associated with human disease (11).

Despite a relatively in-depth knowledge base
for UPEC virulence mechanisms, there is no li-
censed vaccine to prevent UTI, which account
for morbidity worldwide and substantial medi-
cal costs (12). Thus, in order to prevent or con-
trol these infections active clinical and micro-
biological surveillance is essential.

Previous studies showed that vaginal
colonization with *E. coli* is a first step in the
pathogenesis of acute cystitis, and persistent va-
ginal colonization with *E. coli* has been associ-
ated with the development of recurrent urinary
infections in women (13). In this study we se-

---

**Table 2. Virulence scores by phylogenetic group among the vaginal *Escherichia coli* isolates**

<table>
<thead>
<tr>
<th>Phylogenetic group (no. of isolates)</th>
<th>Virulence factors scores</th>
<th>P value&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (19)</td>
<td>4 (1 - 5)</td>
<td>P &lt; 10^{-6}</td>
</tr>
<tr>
<td>B1 (2)</td>
<td>3.5 (3 - 4)</td>
<td>P = 0.0257</td>
</tr>
<tr>
<td>B2 (60)</td>
<td>7 (2 – 9 )</td>
<td>Comparison group</td>
</tr>
<tr>
<td>D (12)</td>
<td>3 (1 – 6)</td>
<td>P &lt; 10^{-6}</td>
</tr>
</tbody>
</table>

<sup>a</sup>The threshold for statistical significance P < 0.05.
lected vaginal *E. coli* strains isolated from women with no UTI symptoms and sought ExPEC-associated traits, which could promote the transition from colonization to infection. Accordingly, we used a molecular approach, namely PCR-based assays, to detect genes encoding putative or proven virulence factors and phylogenetic background as predictors of presumed extraintestinal virulence potential.

ExPEC strains usually belong to phylogenetic group B2 and to a lesser extent to group D, whilst commensal strains are derived from groups A and B1 (14-16). Therefore, the rapid PCR-based phylogenetic typing developed by
Clermont et al. (6) has proven useful for rapidly screening putative ExPEC. According to this method, about two-thirds of the analyzed isolates belonged to group B2. Compared to previous studies examining vaginal *E. coli*, we found a lower percent prevalence of group B2-derived isolates than Obata-Yasuoka et al. (65% vs. 76%), but similar with Watt et al. (68%) and Hilbert et al. (62%). Compared to the data reported by these teams, group D isolates had a lower prevalence in our study (13% vs. 16%, 16%, and 22%, respectively) (17-19), but the commensal phylogenetic group A exhibited a higher prevalence among the isolates investigated by us (20% vs. 8%, 12%, and 8%, respectively). We observed that group B1-derived isolates seemed to be the less suited to colonize the vagina. However, we evaluated the prevalence of the four major phylogenetic groups based on the analysis of a single colony/strain selected from the vaginal swab culture, which could have generated biased results. Nonetheless, we consider that by using this sampling strategy our findings are rather representative for the host’s more abundant or dominant vaginal clone.

By adding the virulence genotyping, we further assessed the pathogenic potential of the microorganisms. According to the ExPEC definition assumed in this study, a significant proportion (81%) of the isolates qualified as ExPEC. The virulence repertoire identified was far from being exhaustive, yet the gene regions targeted could indicate well enough whether the *E. coli* found in the vaginal flora was equipped with a particular set of virulence determinants allowing it to infect the site of temporary residence or the adjacent urinary tract. The isolates examined were genetically diverse, as revealed by the PFGE typing, yet some of them displayed identical virulence genotypes. Among them, 10 isolates comprised genes encoding two adherence systems (type 1 fimbiae and S/F1C fimbiae), two toxins (hemolysin and CNF1) and an iron accumulation system (yersiniabactin) and 12 isolates harbored additionally the P fimbiae-encoding operon. They were all isolates belonging to B2 group. Actually, group B2 strains showed the highest virulence scores and consequently the most diverse virulence profiles (24 of the 35 virulence gene profiles identified). However, ExPEC specific genes were also found in isolates derived from non-B2 groups. The phylogenetic distribution of the targeted virulence genes was obvious, with either the predominance or exclusive concentration of some of them within B2 and/or D groups or the dissemination across all the four groups. This kind of distribution was previously reported (20) being explained by the species’ evolution based on the ongoing vertical inheritance and horizontal transmission of the various virulence factors (21). Of note, within our collection we found a very high prevalence of the *fyuA* and *irp2* genes, which were the second most frequent genes overall. The *fyuA-irp* gene cluster is located on the high-pathogenicity island (HPI) identified primarily in yersiniae and is involved in iron uptake mediated by the siderophore yersiniabactin. The horizontal transfer of the HPI between *Yersinia* spp. and *E. coli* was already documented and its presence was detected in intestinal and extraintestinal clinical *E. coli* isolates (7, 22).

ExPEC strains are known to possess extremely dynamic genomes, which differ considerably with regard to their sizes, the gene contents, and the number of known virulence factors (23). A significant amount of their DNA was obtained through acquisition of foreign DNA from diverse related or non-related donor species by lateral transfer of mobile genetic elements, including pathogenicity islands (PAIs), plasmids, phages, transposons, and insertion elements (24, 25). The simultaneous detection within the same isolate of certain virulence markers were suggestive for the presence of pathogenicity islands sequences (26-28). Sixteen vaginal isolates were positive for *ibeA*, a gene involved in the invasion of brain microvascular endothelial cells, which is localized within the genomic island GimA (29). Forty-three isolates harboured the genes *hlyA* and *cnf1*, char-
acteristic of the PAI IIJ96-like domain and twenty-one of them also had papC sequence and papGIII allele (like strain J96). However, we refrain from making any more comments until doing further investigations to confirm the presence of PAI-associated sequences.

In conclusion, the molecular approach used within our study revealed that most of the vaginal E. coli isolates from otherwise healthy women possessed an assortment of virulence genes which can mediate the steps of extraintestinal infectious process. This implies that an asymptomatic ExPEC reservoir is likely to be present in a substantial percentage of these women. What makes the switch from colonization to infection is still to be understood. However, the presence of these putative pathogenic strains within the host’s own flora might at some point harm the health of woman, her fetus or newborn. According to our knowledge these are the first Romanian published data on the issue. Further studies could contribute to a better evaluation of the prevalence of ExPEC strains in certain groups of women justifying perhaps the need of screening programs similar with those for Streptococcus agalactiae.

Acknowledgements. This work was supported by CNCSIS – UEFISCDI, project number PNII-IDETC21/2008, Romanian Ministry of Education, Research, Youth and Sports.

References


