PapG Alleles of Escherichia coli Strains Causing First-Episode or Recurrent Acute Cystitis in Adult Women

James R. Johnson, Thomas A. Russo, Jennifer J. Brown, and Ann Stapleton

The distribution of the three alleles of the P adhesin gene papG (classes I–III) was assessed among 74 Escherichia coli urine isolates from women with first-episode or recurrent cystitis, and papG genotype was compared with clinical origin, O serogroup, agglutination of Gal(α1-4)Gal—coated latex beads and human or sheep erythrocytes, and hemolysin production. The class-III-only papG genotype (27% of strains) predominated over the I + III (3%) and II-only (7%) genotypes, irrespective of clinical category. In contrast to the class II papG allele, the class III allele was significantly concentrated in serogroups O6 and O18. Agglutination phenotypes corresponded significantly but incompletely with papG genotype, whereas hemolysin production and papG positivity were tightly correlated. These findings suggest that in acute cystitis in adult women, the class III papG allele predominates, confers distinctive agglutination phenotypes, and is restricted to specific E. coli lineages.

P fimbriae, the virulence factor of Escherichia coli most convincingly implicated in the pathogenesis of urinary tract infection (UTI), mediate Gal(α1-4)Gal—specific binding to host surfaces via the adhesin molecule PapG [1]. PapG occurs in three molecular variants (classes I–III) that are encoded by distinct alleles of the adhesin gene papG [2]. The subtly different receptor binding specificities of the three PapG variants may confer differences in host range [3, 4] or capacity for causing different clinical syndromes [5, 6].

The class II papG allele has been associated with pyelonephritis in both children [5] and adult women [6]. In contrast, the class III allele reportedly predominates in cystitis, at least among children [5]; whether this is true also among adult women is unknown. We sought to determine whether the class III papG allele is the predominant papG variant among E. coli isolates from adult women with acute cystitis and whether strains causing first-episode versus recurrent UTI differ with respect to papG allele repertoire. Since the papG alleles may segregate with specific E. coli lineages [2] and confer distinctive agglutination phenotypes [4], and since in some strains hly and pap are genetically linked [1], we also investigated whether, among cystitis isolates from women, papG allele status corresponds with O serogroup, agglutination phenotype, or hemolysin production.

Methods

Pulsed-field gel electrophoresis (PFGE). Strain typing was done by PFGE analysis of macrorestricted genomic DNA, as previously described [7]. E. coli isolates were defined as different strains if from a single subject but exhibiting different genomic patterns or if from different subjects.

Subjects and strains. Seventy-four pretherapy E. coli strains from the urine of 56 adult women (age ≥18 years) that were isolated during 70 episodes of acute uncomplicated cystitis [7, 8] were studied. Subjects either were experiencing their first-ever episode of cystitis (32 episodes in 32 women, 34 total strains) or had recurrent UTI, that is, ≥3 cystitis episodes in the preceding 12 months (38 episodes in 24 women, 40 total strains). All UTI episodes were monomicrobial excepting 4 (2 in first-episode–cystitis and 2 in recurrent-cystitis subjects), each of which involved 2 different E. coli strains.

PapG genotypes. PapG genotypes were determined initially by an allele-specific polymerase chain reaction (PCR) assay [9, 10]. Primers for the class I allele were j96-193f (5’-TCGTGCTCA-GGTCGGAATTT-3’), yielding a 461-bp product; for the class II allele, ia2-383f (5’-GGGATGGCGGGCCCTTGTG-3’) and ia2-572r (5’-CGGGCCCCCAAGTGACTCG-3’), yielding a 190-bp product; and for the class III allele, prs-455f (5’-CCACAAATGC- CATGCCAGAC-3’) and prs-455r (5’-CGGGCCCCCAAGTGACTCG-3’), yielding a 258-bp product [9]. Allele-specific PCR products were resolved by gel electrophoresis. All papG genotypes were confirmed by at least one replicate determination, with discrepancies investigated further (by repeat papG PCR testing, with or without genomic fingerprinting, of multiple isolates of the strain in question) as needed to obtain a reliable result.

For comparison with papG PCR results, strains also were assayed by dot-blot hybridization with DNA probes for each of the three papG alleles and for papC and papEFG. (The latter two probes would be expected to hybridize with any pap-positive strain, irrespective of its specific papG allele configuration, since papC, papE, and papF are highly conserved among different pap operons, independent of their papG content [2].) The three papG allele probes and the papC probe were generated and digoxigenin-labeled (PCR DIG Probe Synthesis kit; Boehringer Mannheim, Indianapolis) using appropriate control strains [9] and primers. For the papG alleles, primers were as described above; for papC, primers were pap1 (5’-GACGGCTGACTCGAGGGTGGTCGGC-3’, forward) and pap2 (5’-ATATCTTCTGACGGATGCATA-
probable explanations for 13 of the discrepancies: contamination including all 27 PCR determinations yielded concordant results for 60 O18 (4/8).

papG -positive vs. -negative, or both positive but for different subsequent analyses. With some positive strains). Only 2 strains exhibited alleles I

papG/papG/papG (papG all, 19% of papG R strains did not differ with respect to the prevalence of ø

Statistical analysis. Individual E. coli strains were used as the unit of analysis throughout. Rough and Ont strains were analyzed as belonging to O groups other than O4, O6, and O18. Comparisons of proportions were tested using Fisher’s exact test. P < .05 was considered significant.

Results

Performance of the papG PCR assay. Initial duplicate papG PCR determinations yielded concordant results for 60 (81%) of the 74 strains and discrepant results (i.e., papG-positive vs. -negative, or both positive but for different papG alleles) for the remaining 14. Subsequent investigation identified probable explanations for 13 of the discrepancies: contamination of a target DNA sample (6), strain substitution (4), either of the above (2), or amplification failure (1). For the single strain whose discrepant results remained unexplained despite the above investigations, the consistent (negative) papG PCR result obtained in 3 of 5 replicate determinations was used in subsequent analyses.

Distribution of papG alleles (table 1). By PCR, 27 (36%) of the 74 strains contained one or more papG alleles. The III-only papG pattern predominated (27% overall, 74% of papG-positive strains). Only 2 strains exhibited alleles I + III (3% overall, 7% of papG-positive strains); 5 had allele II (7% overall, 19% of papG-positive strains) (table 1). For III-only versus I + III or II-only, P = .01 (total population) and P < .001 (papG-positives only). Other possible papG allele configurations (II + III, I + II, I-only) were not encountered. The F and R strains did not differ with respect to the prevalence of papG positivity (table 1).

Dot-blot hybridization with the three papG allele-specific DNA probes gave results identical to those of the papG PCR assay for 72 (97%) of the 74 strains. The 2 strains that gave discrepant results were investigated further. One had been identified as papG-negative by PCR (3 of 5 runs with no PCR product, 2 runs with papG allele III). Probing of multiple colonies yielded both positive and negative results for papC, papEFG, and papG (allele III), with consistent results across probes for an individual colony but discrepant results between colonies. By papG PCR, the different colonies yielded results consistent with their probe results. The second discrepant strain, which by PCR was consistently positive for papG allele II only, was reproducibly positive by DNA probe for both allele II and (weakly) allele III. This strain was papC and papEFG probe-positive. In no instance did the papG PCR assay detect a papG allele that was not confirmed by DNA probe. In subsequent analyses, papG genotypes as defined by PCR were used for all 74 strains.

O serogroup versus papG alleles (table 1). The 69 strains that were O-typeable exhibited 27 different O antigens. Four strains were Ont and one was rough. The 6 most frequent O groups, each represented by ≥ 3 strains, accounted for 54% of strains and included groups O1 (5 strains), O2 (4), O4 (3), O6 (14), O18 (11), and O75 (3).

papG positivity was concentrated in groups O4, O6, and O18 (P < .001, groups O4/O6/O18 [22/28] versus other strains [5/46]) (table 1). The I + III papG allele configuration was encountered only in serogroup O4. The III-only pattern occurred predominantly in serogroups O6 and O18 (P < .001, groups O6/O18 [18/25] versus other strains [2/49]). In contrast, the proportion of papG-positive strains exhibiting the II-only pattern was significantly lower among O6 and O18 strains than among other strains (P = .02, O6/O18 [1/25] versus non-O6/O18 [4/8]).

Agglutination phenotype versus papG alleles (table 1). Overall, 40 (54%) of the 74 strains agglutinated some combination of P beads and human (H) or sheep (S) erythrocytes, including all 27 papG-positive strains but only 13 (21%) of 47 papG-negative strains (P < .001) (table 2). P bead agglutination was sensitive (7/7) but not specific (7/10) for I + III or II-only papG status (table 1). Agglutination of sheep erythrocytes occurred with all papG-positive strains (100% sensitivity) but was observed with strains of all papG allele patterns and even with some papG-negative strains (table 1). Agglutination of human erythrocytes was common (36/40) among agglutinating strains, irrespective of papG status.

The P’S’H’+ agglutination pattern was confined to papG-positive strains and, among papG-positive strains, was specific to the I + III or II-only strains (P < .001, I + III or II-only versus III-only or versus all other strains). In contrast, the P’S’H’ pattern characterized the III-only papG-positive strains (P < .001, III-only versus other papG-positive strains or versus all other strains). However, 4 (20%) of the III-only strains exhibited an alternative pattern, which always included sheep erythrocytes (table 2). Furthermore, the P’S’H’ pattern
occurred with 3 non–III-only strains, all of which were \( \text{papG} \)-negative by both PCR and DNA probe hybridization (table 1). Two of the 3 were probe-positive for \( \text{papEFG} \); 1 of these (serogroup O1) was probe-positive and another (serogroup O21) probe-negative for \( \text{papC} \). The third strain (serogroup O2) was probe-negative for both \( \text{papEFG} \) and \( \text{papC} \).

The \( P^+S^+H^+ \) (NPMR) pattern was the most common agglutination pattern of the \( \text{papG} \)-negative strains and was specific to these strains \( (P < .001 \) vs. \( \text{papG} \)-positive strains). Of the NPMR strains, 4 (O9:Hnt, O1:Hnt, O46:H31, and O15,O143:NM) were probe-positive for \( \text{papEFG} \) but negative for \( \text{papG} \). Of these, 2 (O9 and O1) were positive or variably positive for \( \text{papC} \), whereas 2 (O46 and O15,O143) were \( \text{papC} \)-negative. The other 4 NPMR strains (O75:NM \( \times 2 \), O81:H7,H24, and O25:Hnt) were uniformly negative for \( \text{papEFG} \), \( \text{papC} \), and \( \text{papG} \) (the latter by both probe and PCR).

Atypical agglutination patterns that were observed with only 1 strain each included \( P^+S^-H^+ \) (O8:NM) and \( P^+S^-H^- \) (O95:NM). The former strain was probe-positive and the latter, probe-negative for both \( \text{papC} \) and \( \text{papEFG} \). Both strains were negative for \( \text{papG} \) by both probe and PCR.

**Hemolysin production.** Hemolysin production corresponded closely with \( \text{papG} \) positivity: 24 of 27 \( \text{papG} \)-positive strains were hemolytic versus 1 of 47 \( \text{papG} \)-negative strains \( (P < .001) \). Discrepancies were confined to serogroups O2 and O18 \( (P = .001 \) vs. other strains), in that 2 III-only O18 strains and 1 II-only O2 strain were nonhemolytic but \( \text{papG} \)-positive, and a single O2 strain (which was the \( \text{papEFG} \)- and \( \text{papC} \)-negative strain that exhibited a class III–like \( P^+S^-H^- \) agglutination pattern) was hemolytic but \( \text{papG} \)-negative. All 4 strains that were nonconcordant for hemolysin versus \( \text{papG} \) were from recurrent-cystitis subjects \( (P < .001, \) recurrent vs. first-episode cystitis).

**Discussion**

We found that among *E. coli* strains from adult women with first-episode or recurrent acute cystitis, the class III \( \text{papG} \) allele

---

**Table 1.** Distribution of \( \text{papG} \) alleles by clinical category, O serogroup, and agglutination pattern among 74 *E. coli* strains from women with acute cystitis.

<table>
<thead>
<tr>
<th>( \text{papG} ) PCR result</th>
<th>First-episode (F) or recurrent (R) cystitis</th>
<th>O serogroup</th>
<th>Agglutination pattern for ( P ) beads (P), sheep erythrocytes (S), and human erythrocytes (H)(^*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{papG} ) PCR result</td>
<td>Total ((n = 74))</td>
<td>F ((n = 34))</td>
<td>R ((n = 40))</td>
</tr>
<tr>
<td>I + III</td>
<td>2 (3)</td>
<td>0 (0)</td>
<td>2 (5)</td>
</tr>
<tr>
<td>III only(^*)</td>
<td>20 (27)</td>
<td>11 (32)</td>
<td>9 (23)</td>
</tr>
<tr>
<td>II only(^*)</td>
<td>5 (7)</td>
<td>2 (6)</td>
<td>3 (8)</td>
</tr>
<tr>
<td>Any ( \text{papG} )</td>
<td>27 (36)</td>
<td>13 (38)</td>
<td>14 (35)</td>
</tr>
<tr>
<td>No ( \text{papG} )</td>
<td>47 (64)</td>
<td>21 (62)</td>
<td>26 (65)</td>
</tr>
</tbody>
</table>

\(^*\) Includes O-nontypeable and rough strains.

\(^2\) \( P^+ \), \( S^- \), and \( H^- \) indicate presence and \( P^- \), \( S^+ \), and \( H^+ \) indicate absence of agglutination.

\(^1\) One strain (class II–positive by both polymerase chain reaction (PCR) and DNA probe; listed above as II only) was weakly positive also for class III by probe but not PCR.

---

**Table 2.** Distribution of \( \text{papG} \) alleles according to clinical category (from literature and present study).

<table>
<thead>
<tr>
<th>( \text{papG} ) alleles</th>
<th>Pyelonephritis</th>
<th>Cystitis</th>
<th>ABU</th>
<th>Fecal</th>
<th>Pyelonephritis in women*</th>
<th>Cystitis in women(^1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>((n = 67))</td>
<td>((n = 71))</td>
<td>((n = 66))</td>
<td>((n = 74))</td>
<td>((n = 60))</td>
<td>((n = 15))</td>
</tr>
<tr>
<td>I + III</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>III only</td>
<td>4 (6)</td>
<td>23 (32)</td>
<td>8 (12)</td>
<td>3 (4)</td>
<td>2 (3)</td>
<td>0</td>
</tr>
<tr>
<td>II + III</td>
<td>7 (10)</td>
<td>3 (4)</td>
<td>1 (2)</td>
<td>1 (1)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>II only</td>
<td>41 (61)</td>
<td>15 (21)</td>
<td>19 (29)</td>
<td>12 (16)</td>
<td>44 (73)</td>
<td>14 (93)</td>
</tr>
</tbody>
</table>

NOTE. ABU, asymptomatic bacteriuria.

\( \text{papG} \) alleles determined by *DNA probe hybridization or \( \)polymerase chain reaction assays.
was by far the predominant \( \text{papG} \) variant. This is consistent with a previous report that the class III allele predominates among cystitis isolates from children (37% of strains vs. 25% for the class II allele) [5]. These associations of the class III \( \text{papG} \) allele with cystitis contrast with the class II allele’s reported associations with pyelonephritis among children [5] and adult women (particularly those with bacteremia) [6] (table 2) and suggest that \( \text{papG} \) alleles segregate according to “level” of infection in the urinary tract, with the class III allele participating to a greater extent in bladder infection and the class II allele in kidney or bloodstream infection [5].

However, the frequency of the class II allele was substantially higher among pediatric cystitis isolates [5] than observed here among adult cystitis isolates (table 2). Similarly, data from one of our laboratories demonstrate a striking predominance of the class II \( \text{papG} \) allele among pediatric cystitis isolates from Cleveland (Johnson JR, unpublished data) and a substantial prevalence of the class III allele among mixed-source bacteremia isolates from the United States and Kenya [11a]. Thus, associations of the \( \text{papG} \) alleles with specific clinical syndromes may depend on the particular population studied.

The observed predominance of the class III \( \text{papG} \) allele provides further evidence that this adhesin is of pathogenic importance, not only for dogs [4] but also for humans. This would be expected in view of the class III adhesin’s recognition of a receptor that is widely prevalent among humans, that is, stage-specific antigen 4 (SSEA-4; also LKE antigen), and is expressed within the human urinary tract [12]. One potential clinical implication of this finding is that the class III \( \text{PapG} \) variant may need to be included in future adhesin-based anti-UTI vaccines [13] if their goal is to prevent cystitis in women or children.

The similarity of strains from first-episode and recurrent cystitis with respect to \( \text{papG} \) allele repertoires (table 1) is consistent with such strains’ similar profiles for other virulence properties [8]. It suggests that whatever the factors are that determine the increased tendency to UTI among women with recurrent UTI, they do not discriminate between the \( \text{PapG} \) variants. This is consistent with the observation that the Gal(\( \alpha1-4 \))Gal–containing glycolipids (including SSEA-4) that are uniquely present on vaginal epithelial cells of (UTI-prone) nonsecretor women [14] are effective receptors for all three \( \text{PapG} \) variants [15], so presumably would not select for one \( \text{PapG} \) variant over another. To the extent that recurrent UTI is associated with nonsecretor status [1, 14], the similar distribution of \( \text{papG} \) alleles among first-episode and recurrent strains also argues that class III–positive strains do not specifically target A\(_S\) secretor hosts, as has been proposed [3, 16].

The identification of 2 strains with the \( \text{I} \) or III \( \text{papG} \) allele configuration (table 1) demonstrates that this genotype, although uncommon among clinical isolates (table 2), is not confined to source strain J96 as was previously thought [5, 16]. Other evidence indicates that this genotype is characteristic of a clonal group within serogroup O4 whose members (including J96) have caused cystitis, pyelonephritis, urosepsis, and bacteremia of unknown source in both the United States [10] and Europe (Johnson JR, unpublished data).

The class III \( \text{papG} \) allele was significantly concentrated in the two most prevalent serogroups (O6, O18) and the class II allele in other serogroups (table 1). To our knowledge, this is the first report of serogroup-specific associations of \( \text{papG} \) alleles. Our findings are consistent with the observation that the \( \text{papG} \) alleles segregate according to \( \text{E. coli} \) lineage in a fashion suggesting both vertical and horizontal transmission [2]. We have observed similar serogroup-specific associations of the \( \text{papG} \) alleles in other collections of clinical \( \text{E. coli} \) isolates [11a], evidence that they occur broadly.

The degree of phenotypic diversity noted among the III-only strains was much lower than in the only previous large-scale evaluation of phenotype versus \( \text{papG} \) genotype among human clinical isolates [5]. We found that differentiation of III-only strains from I + III or II-only strains by agglutination phenotype required P beads, since almost all \( \text{papG} \)-positive strains (irrespective of \( \text{papG} \) allele repertoire) agglutinated both human and sheep erythrocytes (table 2). Unfortunately, P beads are no longer commercially available. HA assays using human, rabbit, and sheep erythrocytes have been proposed as a means for distinguishing among the three \( \text{PapG} \) variants without the use of receptor-coated beads [4]. However, slide HA assays are highly irreproducible and do not reliably resolve the \( \text{PapG} \) variants [17]. Microtiter tray HA assays, incorporating diverse erythrocyte types, may provide a more reproducible and discriminating method for assessing \( \text{papG} \)-specific agglutination phenotypes when \( P \) beads are not available (Johnson JR, unpublished data).

Three PCR \( \text{papG} \)-negative strains in the present study gave the same P’S’H’ agglutination pattern as did most class III \( \text{papG} \)-positive strains (table 2). That these 3 strains also failed to hybridize with DNA probes for the three \( \text{papG} \) alleles confirms that they do not contain sequences closely related to those of the recognized variants of \( \text{papG} \). Another indication that these strains do not contain \( \text{papG} \) allele III is that they were from serogroups not otherwise associated with this allele. However, 2 of these strains were probe-positive for \( \text{papEFG} \), and 1 of these 2 was probe-positive also for \( \text{papC} \), indicative of the presence of at least partial copies of \( \text{pap} \) (or degenerate \( \text{pap} \) sequences). Another \( \text{papG} \)-negative strain, which gave an unusual P’S’H’ agglutination pattern, was positive for both \( \text{papC} \) and \( \text{papEFG} \). Whether, in these enigmatic strains, HA is mediated by an as-yet-unrecognized \( \text{papG} \) variant (the possible existence of which is supported also by previous reports of phenotypic Pap positivity among \( \text{papG} \)-probe-negative strains [5, 6]), or instead is due to some alternative non-\( \text{pap} \) (i.e., NPMR) adhesin, remains to be determined.

Results from the \( \text{papG} \) PCR assay corresponded closely with those from dot-blot hybridization using allele-specific DNA probes. In no instance did the PCR assay give a positive result that was not confirmed by probe, and in only 2 instances did probing suggest the presence of \( \text{papG} \) sequences that were not consistently detected by PCR. For each of these strains, the
positive probe results were either irreproducible or weak. The strain with variable probe results also had variable papG PCR results, which actually corresponded (at the individual colony level) with the probing. The weak class III probe result for the strain that had only allele II by PCR suggested the possibility of substantial target sequence divergence of this strain’s version of papG allele III away from the archetypal representatives of allele III on which the PCR primers and DNA probe were based. Even if this strain were to be considered to have allele III (along with its clear-cut allele II), this would only strengthen our conclusions regarding the predominance of allele III in the study population. The observed occasional irreproducibility of the papG PCR assay, for diverse causes, indicates the importance of performing replicate determinations using different target DNA samples when testing clinical collections with the PCR assay.

In summary, we found that among E. coli isolates from women with first-episode or recurrent acute cystitis, the class III allele was the predominant papG variant and was concentrated in serogroups O6 and O18. Agglutination phenotypes corresponded moderately well with papG genotypes. Hemolysin production and papG positivity were tightly linked. These findings suggest that the adherence phenotypes conferred by the class III papG allele may be well suited for the pathogenesis of cystitis in adult women, that the class III variant of PapG should be included in future adhesin-specific vaccines against cystitis, and that certain E. coli lineages may selectively participate in specific clinical syndromes because of their associated virulence traits.

Acknowledgments

Parvia Ahmed assisted in the laboratory. Mary Hayes and Ann Emery helped prepare the manuscript.

References