

A Novel 7-Single Nucleotide Polymorphism-Based Clonotyping Test Allows Rapid Prediction of Antimicrobial Susceptibility of Extraintestinal *Escherichia coli* Directly From Urine Specimens

Veronika Tchesnokova,¹ Hovhannes Avagyan,¹ Mariya Billig,¹ Sujay Chattopadhyay,¹ Pavel Aprikian,³ Diana Chan,¹ Julietta Pseunova,⁴ Elena Rechkina,³ Kim Riddell,⁵ Delia Scholes,⁵ Ferric C. Fang,^{1,2} James R. Johnson,⁶ and Evgeni V. Sokurenko¹

Departments of ¹Microbiology, ²Laboratory Medicine, University of Washington School of Medicine, Seattle, and ³ID Genomics, Inc., Seattle, Washington; ⁴Gemotest, Moscow, Russia; ⁵GroupHealth Cooperative, Seattle, Washington; ⁶VA Medical Center and University of Minnesota, Minneapolis

Background. *Escherichia coli* is a highly clonal pathogen. Extraintestinal isolates belong to a limited number of genetically related groups, which often exhibit characteristic antimicrobial resistance profiles.

Methods. We developed a rapid clonotyping method for extraintestinal *E coli* based on detection of the presence or absence of 7 single nucleotide polymorphisms (SNPs) within 2 genes (*fumC* and *fimH*). A reference set of 2559 *E coli* isolates, primarily of urinary origin, was used to predict the resolving power of the 7-SNP-based typing method, and 582 representative strains from this set were used to evaluate test robustness.

Results. Fifty-four unique SNP combinations (“septatypes”) were identified in the reference strains. These septatypes yielded a clonal group resolution power on par with that of traditional multilocus sequence typing. In 72% of isolates, septatype identity predicted sequence type identity with at least 90% (mean, 97%) accuracy. Most septatypes exhibited highly distinctive antimicrobial susceptibility profiles. The 7-SNP-based test could be performed with high specificity and sensitivity using single or multiplex conventional polymerase chain reaction (PCR) and quantitative PCR. In the latter format, *E coli* presence and septatype identity were determined directly in urine specimens within 45 minutes with bacterial loads as low as 10² colony-forming units/mL and, at clinically significant bacterial loads, with 100% sensitivity and specificity.

Conclusion. 7-SNP-based typing of *E coli* can be used for both epidemiological studies and clinical diagnostics, which could greatly improve the empirical selection of antimicrobial therapy.

Keywords. antibiotic resistance; clonal typing; *Escherichia coli*; personal therapy; point-of-care testing.

Escherichia coli is a major causative agent of acute infections, both gastrointestinal and extraintestinal [1, 2]. Extraintestinal pathogenic *E coli* (ExPEC), a subset of *E coli*, is the major cause of urinary tract infections (UTIs), newborn meningitis, and bacterial sepsis. Extraintestinal pathogenic *E coli* strains are highly clonal, with certain closely related clonal groups associated with characteristic antimicrobial resistance profiles [3–6]. Therefore, subspecies delineation of ExPEC clonal groups has significant potential for improving the diagnostics of antimicrobial resistance, especially if it can be performed directly from the clinical specimen and quickly (eg, <1 hour).

The current standard for clonal typing is multilocus sequence typing (MLST) [7], which involves sequencing of approximately 7 housekeeping genes and splitting *E coli* into sequence types (STs). However, MLST is laborious and time consuming, and thus it is not well suited for clinical or epidemiological applications. We have recently introduced a method for predicting clonal types and associated resistance profiles based on sequencing of only 2 genes: *fumC* and *fimH* (ie, CH typing). This scheme was shown to subdivide *E coli* into a large number of clonal groups or combinations thereof (clonotypes), and it was effective both in predicting and subdividing specific ExPEC STs [3, 8]. We have successfully used CH typing as a proof-of-principle for the use of clonotyping to predict antibiotic resistance among clinical isolates. However, the *fumC/fimH* clonotyping scheme still involves DNA sequencing, limiting its broader application.

In this study, we report a novel method for ExPEC clonotyping that incorporates polymerase chain reaction (PCR) or quantitative PCR (qPCR)-based determination of the presence or absence of just 7 single nucleotide polymorphisms (SNPs) within *fumC* and *fimH*. We demonstrate that alternative binary

Received 25 December 2015; accepted 2 January 2016.

Correspondence: E. V. Sokurenko, Department of Microbiology, Health Sciences Building, Room J267a, 1705 NE Pacific St, Seattle, WA 98195-7735 (evs@u.washington.edu).

Open Forum Infectious Diseases®

© The Author 2016. Published by Oxford University Press on behalf of the Infectious Diseases Society of America. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact journals.permissions@oup.com. DOI: 10.1093/ofid/ofw002

combinations of the 7 SNPs splits ExPEC clinical isolates into several dozen clonal types that exhibit distinct antibiotic susceptibility profiles. We show that the 7-SNP typing test can rapidly identify clonal types of *E. coli* directly from urine specimens, demonstrating the potential to predict antibiotic resistance using a clonal diagnostics approach in the point-of-care setting.

MATERIALS AND METHODS

Bacterial Isolates

The reference set consisted of 2559 random single-patient clinical *E. coli* isolates of primarily extraintestinal origin, recovered between October 2010 and June 2013 at 8 clinical microbiology laboratories. The source laboratories serve distinct patient populations in Seattle, Washington (Group Health Cooperative, King County Harborview Medical Center, Seattle Children's Hospital, University of Washington Medical Center), Minneapolis, Minnesota (Veterans Affairs Medical Center), Muenster, Germany (University Hospital), Moscow, Russia (Gemotest Center), and Wrocław, Poland (Medical University of Wrocław). The major source of isolates was urine (67.1%), followed by feces (6.8%) and blood (4.6%), with the remaining isolates originating from wounds, the abdomen, and other extraintestinal compartments. Non-*E. coli* species of *Enterobacteriaceae* were isolated in our laboratory during routine urine and fecal screening and characterized by standard microbiological techniques and 16S sequencing.

Clonal Typing of *Escherichia coli* Isolates Using Multilocus Sequence Typing and *fumC* and *fimH* Typing

Conventional MLST was performed either in full using the standard MLST scheme of 7 loci (*adk*, *fumC*, *gyrB*, *icd*, *mdh*, *purA*, and *recA*) [4] or as a combination of partial MLST and *fumC*/*fimH* typing, as described previously [3]. For simplicity of presentation, some ST complexes, as identified by the eBURST version 3 software, were defined operationally as a single ST and were designated based on the ST identified as the cluster's founder (see [Supplementary Table 3](#)) [9]. Such complexes accounted collectively for 17.5% of isolates, with most comprised of minor groups.

Development of a 7-Single Nucleotide Polymorphism-Based Method of *Escherichia coli* Typing

fumC and *fimH* were selected as the gene targets for development of the SNP-based *E. coli* clonotyping method because of the previously demonstrated power of *fumC*/*fimH* clonotyping to predict ST-based clonal groups and to subdivide them into smaller subclones [3]. To guide SNP selection, a proprietary algorithm developed by ID Genomics, Inc. (Seattle, WA) was used to select 7 SNPs that included 3 SNPs from *fumC* (positions 63, 248, and 380) and 4 from *fimH* (positions 108, 162, 233, and 483). The 7-SNP format allowed the PCR test to be performed using an 8-tube/8-well configuration, with the 8-hour tube or well used for a control.

To record SNP presence or absence, we used a numerical key algorithm, as in other binary coding systems [10]. For this, the 7 SNPs were split into 3 groups. The first group included 3 SNPs (*fumC*-63, *fumC*-248, *fumC*-380), the second group another 3 SNPs (*fimH*-108, *fimH*-162, *fimH*-233), and the third group the remaining SNP (*fimH*-483). Presence of the first SNP within a group was scored as 1, presence of the second SNP was scored as 2, and presence of the third SNP was scored as 4, whereas SNP absence was scored as 0. Each isolate was assigned a score based on the sum of the scores for each SNP group ([Supplementary Table 1](#)). For example, the 101-011-1 binary combination (where "1" and "0" is SNP presence and absence, respectively) was recorded as "561", 010-111-0 as "270", etc. We designated this typing method as the "7-SNP typing test" and the corresponding clonotypes as "septa-types".

Typing of *Escherichia coli* Isolates Using the 7-Single Nucleotide Polymorphism Typing Test

For the 7-SNP typing test, 7 primer pairs were designed and optimized to detect the SNPs, with *uidA* locus-specific primers added as an *E. coli* positive control ([Supplementary Table 2](#)). The performance of the primers was tested using 582 isolates that contained, collectively, all variant *fumC* and *fimH* alleles identified in our reference set of 2559 *E. coli* that had any polymorphisms in the primer-binding region ([Supplementary Table 2](#)).

Testing was done in 3 different formats: singleplex or multiplex conventional PCR, or singleplex qPCR. Singleplex 7-SNP typing consisted of 8 independent reactions using the above-described primers. Multiplex 7-SNP typing used some of the primer pairs combined into 3 total reactions, as shown in [Supplementary Table 2](#). The first triplex reaction contained 3 SNP-specific forward primers (*fumC*-63, *fumC*-248, and *fimH*-483) and 2 common reverse primers (1 each for *fumC* and *fimH*). The second triplex reaction contained the *uidA* primers, 2 SNP-specific forward primers (*fimH*-108 and *fimH*-162), and a common reverse primer for *fimH*. The duplex reaction contained the same 2 primer pairs as used in singleplex PCR (for *fumC*-380 and *fimH*-233).

Of the 582 isolates representing the range of *fumC*/*fimH* allele diversity that were used to test primer performance, 310 underwent full 7-SNP typing test (including the *uidA* control), which was done using a single format for 180 isolates, 2 formats for 121 isolates, and all 3 formats for 24 isolates. The remaining 272 test isolates were screened using primers for individual SNPs. If the 7-SNP test failed to detect the correct septatype of a particular *fimH* or *fumC* allele, the test was repeated to confirm or refute this failure.

Test reactions used JumpStart PCR master mix (Sigma) and the following conditions: 2 minutes initial denaturation at 95°C, followed by 27 cycles of 15 seconds at 95°C, 15 seconds at 57°C,

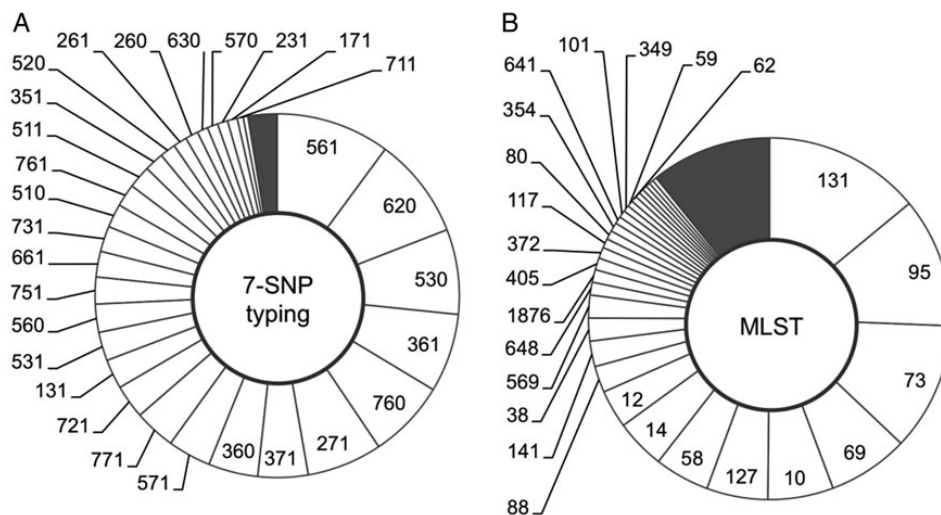


Figure 1. (A) Comparison of the diversity detected by 7-single nucleotide polymorphism (SNP) typing or (B) conventional multilocus sequence typing (MLST). Reference set of 2599 clinical *Escherichia coli* was split into clonotypes using either 7-SNP typing or conventional MLST (A). The segments of the doughnuts represent individual clonotypes; their size reflects their prevalence within the population. Clonotypes are sorted in descending order of prevalence. All nonminor clonotypes (>0.5% of population) are labeled.

and 30 seconds at 72°C. Amplification products were analyzed by 2% agarose gel electrophoresis. The qPCR 7-SNP typing test was performed by real-time quantitative PCR on a Rotor-Gene Q MDx instrument (QIAGEN) using the SYBR-Green PCR Kit (QIAGEN). Unless stated otherwise, the qPCR reaction conditions were as follows: 3 minutes at 95°C, followed by 30 cycles of 5 seconds at 95°C, 10 seconds at 57°C, and 10 seconds at 72°C, with signal acquisition at the elongation step. As in the singleplex test, the qPCR 7-SNP typing test consisted of 8 independent reactions.

7-Single Nucleotide Polymorphism Typing of *Escherichia coli* From Urine Samples

To validate the performance of the 7-SNP typing test, 160 random urine samples were obtained from the clinical microbiology laboratory at the Harborview Medical Center (Seattle, WA). Each sample was processed as follows: 1 mL urine was added to 50 µL of a 20% suspension of Chelex (Bio-Rad) in sterile water and centrifuged 1 minutes at 12 000 rpm. The pellet with Chelex was resuspended in 100 µL sterile water, heated for 5 minutes at 96°C, and centrifuged again, with the supernatant used as template for the qPCR-based 7-SNP typing test. The load of *E. coli* in urine was determined from the *uidA* threshold cycle based on a standard calibration curve. (Sterile urine samples yielded no *uidA* signal.) In parallel, 10 µL urine were plated on McConkey agar to detect the growth of *E. coli*. Cultured *E. coli* isolates underwent clonotyping as a control.

Antibacterial Susceptibility

Susceptibility to 7 antibiotics was determined using a standardized disk diffusion method according to Clinical and Laboratory Standards Institute guidelines [11].

Statistical Analysis

Simpson's diversity index was calculated using the formula: $D = 1 / \sum [n(n-1) / N(N-1)]$, where n is the number of *E. coli* isolates in a particular clonotype, and N is the total number of isolates. Comparisons of proportions were tested using a 2-tailed Fisher's exact test. Bacterial loads detected in urine by the 7-SNP typing test versus culture were compared using the 2-sided paired t test.

RESULTS

7-Single Nucleotide Polymorphism Typing Provides High Resolution of *Escherichia coli* Clonotypes

The 7-SNP typing test was evaluated by computer-based analysis using the *fumC* and *fimH* nucleotide sequences of 2559 recent extraintestinal isolates from 8 clinical microbiology laboratories in the United States, Germany, Poland, and Russia. According to the analysis, the 2559 isolates were divided among 54 unique binary septatypes (Figure 1A). Six "major" septatypes included more than 5% of isolates and comprised 47% of all isolates. Twenty-three "intermediate" septatypes included 0.5%–5% of isolates each and together comprised 50% of the isolates. The remaining 25 septatypes were "minor", each including <0.5% of isolates. The overall diversity of major and intermediate septatypes was similar to that of MLST (Figure 1B), which identified 6 major and 19 intermediate STs of the same size as defined by 7-SNP typing. Although there were nearly 3 times as many STs as septatypes (175 vs 54), this was due entirely to the greater number of minor STs (150 in total) that together comprised only 13% of the isolates. This parity of two typing schemes is reflected in comparable Simpson's diversity index values for 7-SNP typing and MLST (0.949 and 0.934, respectively).

Overall, in 27 septatypes (4 major, 10 intermediate, 15 minor) >90% of isolates (97.2% on average) belonged to a single

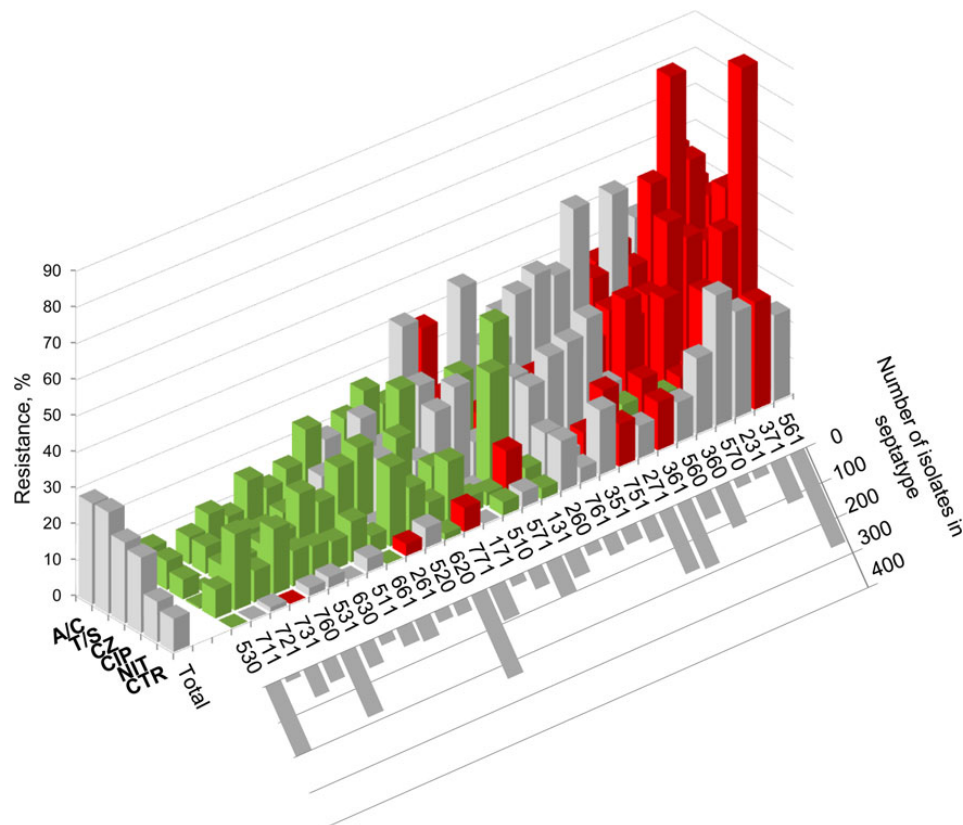


Figure 2. Clonotype-specific antibacterial resistance profile. The reference set of 2599 *Escherichia coli* isolates was split into septatypes by computer sequence analysis. The prevalence of resistance within individual nonminor septatypes (>0.5% of isolates each) to amoxicillin/clavulanate (A/C), trimethoprim/sulfamethoxazole (T/S), cefazolin (CZ), ciprofloxacin (CIP), nitrofurantoin (NIT), and ceftriaxone (CTR) is plotted as vertical columns. Columns are color coded to indicate whether, compared with the total population, the prevalence of resistance in this clonotype is significantly ($P < .05$) higher (red) or lower (green) or is not significantly different (gray). The graph inserted to the lower right of the main graph shows the number of isolates in individual septatypes.

ST (Supplementary Table 3). Such clonally homogeneous septatypes accounted for 54.1% of all isolates. Overall, 17 STs could be predicted by the homogeneous septatypes, including multidrug-resistant ST131, ST69, and ST38, and relatively susceptible ST95, ST73, ST14, ST12, ST569, and ST117. Moreover, several major STs, such as ST131, ST95, and ST73, were split by 7-SNP typing into smaller sub-ST clonal groups. For example, ST131 was split into its *fimH*-based H30, H41, and H22 subclones [12], ST95 into its H41, H15, and H30 subclones, and ST73 into its H9, H10, and H30 subclones [13].

The remaining 27 septatypes were clonally heterogeneous, ie, <90% of each septatype were of the same ST. Such major STs as drug-resistant ST58, ST88, ST354, ST648, and susceptible ST127, ST141, ST10, ST1876, were either distributed among different septatypes or could not be identified with high confidence.

Thus, binary combinations of 7 SNPs selected from 2 highly variable genes, *fimH* and *fumC*, can genotype *E coli* isolates with high resolution by splitting them into a large number of clonotypes. In most isolates, 7-SNP typing accurately predicted the ST of origin and, often, also the sub-ST.

7-Single Nucleotide Polymorphism Typing Sorts Isolates Into Groups With Distinctive Antibacterial Susceptibility Profiles

In the total *E coli* study population, the prevalence of resistance to 6 antibiotics commonly used for empirical treatment of UTI was as follows: amoxicillin/clavulanate and trimethoprim/sulfamethoxazole, 28% each; cefazolin, 23%; ciprofloxacin, 21%; nitrofurantoin, 11%; and ceftriaxone, 9%. We compared resistance prevalence in the total population to that within individual septatypes.

For each major and intermediate septatype, the prevalence of resistance to at least 1 antibiotic was significantly lower (“susceptible” septatype) or higher (“resistant” septatype) than in the total population (Figure 2). More importantly, 7-SNP typing split several prominent multidrug-resistant STs into smaller clonal groups with distinct antimicrobial resistance profiles in comparison with the total population and the corresponding ST (Table 1). For example, ST131, the predominant multidrug-resistant *E coli* clonal group, was split into 3 septatypes. Septatype 561 (corresponding to the ST131-H30 subclone) had a high prevalence of ciprofloxacin resistance (90.6%), whereas septatypes 560 (ST131-H41) and 510 (ST131-H22)

Table 1. Splitting of Multilocus Sequence Typing-Based Sequence Types Into Septatypes With Distinct Antimicrobial Profiles^a

ST	Septatype	No. Isolates	Resistance Prevalence (%)					
			A/C	CZ	CTR	T/S	CIP	NIT
ST131	Total	352	48	46	20	46	63	<i>10</i>
	561	242	52	54	26	47	91	<i>13</i>
	560	60	<i>38</i>	<i>25</i>	<i>10</i>	58	2	<i>7</i>
	510	49	43	<i>31</i>	<i>4</i>	<i>25</i>	2	2
ST69	Total	188	<i>26</i>	<i>23</i>	<i>8</i>	59	<i>20</i>	<i>12</i>
	271	161	<i>26</i>	<i>22</i>	<i>7</i>	62	<i>15</i>	<i>13</i>
	351	13	<i>39</i>	<i>23</i>	<i>23</i>	69	92	<i>8</i>
	261	9	0	<i>11</i>	<i>0</i>	0	0	<i>11</i>
ST58	Total	127	<i>33</i>	<i>18</i>	<i>13</i>	48	<i>27</i>	19
	361	48	<i>21</i>	<i>15</i>	<i>9</i>	<i>38</i>	<i>28</i>	<i>13</i>
	360	39	44	<i>28</i>	<i>21</i>	46	51	<i>15</i>
	371	27	48	<i>15</i>	<i>11</i>	78	4	44

Abbreviations: A/C, amoxicillin/clavulanate; CIP, ciprofloxacin; CTR, ceftriaxone; CZ, cefazolin; MLST, multilocus sequence typing; NIT, nitrofurantoin; ST, sequence type; T/S, trimethoprim/sulfamethoxazole.

^a The prevalence of resistance within each ST and septatype is compared with that within the total *Escherichia coli* population by using a 2-sided exact Fisher's test. Bold black font indicates $P < .05$ vs average; italic font indicates no significant difference from the population average. "Total" denotes all isolates that belong to the corresponding ST.

were almost entirely susceptible. In contrast, both 561 and 560, but to a much lesser extent 510, had a high prevalence of trimethoprim/sulfamethoxazole resistance. In ST69 (aka clonal group A), notorious for trimethoprim/sulfamethoxazole resistance, the 7-SNP test resolved septatype 351 (typically ciprofloxacin-resistant) and septatype 261 (typically pan-susceptible, including to trimethoprim/sulfamethoxazole. The 7-SNP test yielded similar resistance-informative splits within ST58.

Overall, the by-antibiotic proportion of isolates belonging to a resistant or susceptible septatype was greatest for ciprofloxacin, for which 92% of isolates belonged to either resistant (34%) or susceptible (59%) septatypes (Figure 3), with the average resistance prevalence in the former and latter groups being 51% vs 4.4%, respectively. For other antibiotics, such differences were not as pronounced, but for all antibiotics except nitrofurantoin, most isolates were in septatypes with a resistance profile distinct from the population average. It is notable that for all antibiotics except amoxicillin/clavulanate, the average resistance prevalence within susceptible septatypes was <10%.

Overall, 7-SNP typing performed well in subdividing *E. coli* clinical isolates into clonotypes with distinctive resistance profiles. Thus, it has the potential to provide much greater accuracy than a species-wide antibiogram in predicting the susceptibility of individual isolates.

Experimental Validation of the 7-Single Nucleotide Polymorphism Typing Test

We designed 7 pairs of SNP-specific primers and a pair of *uidA*-specific primers (for detection of *E. coli*) (Supplementary Table 2) that, in conventional or qPCR, yielded the predicted

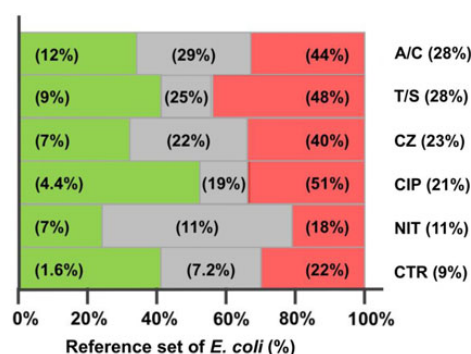


Figure 3. 7-single nucleotide polymorphism (SNP) typing-based detection of *Escherichia coli* taxa with divergent antimicrobial resistance phenotypes. The reference set of 2599 *E. coli* isolates was analyzed as shown in Figure 2, namely, 7-SNP typing was used to split the set into individual clonotypes, for which the prevalence of resistance to 6 antibiotics was calculated. Each bar represents the total reference set analyzed by 7-SNP typing. Each bar is split into 3 zones, for isolates belonging to clonotypes for which the resistance prevalence to a given antibiotic is significantly below (green), significantly above (red), or indistinguishable from (gray) the overall population average. Percentage values within a colored zone denote the resistance prevalence within that subset. Antibiotics are listed to right of the graph, with their overall resistance prevalence (Abbreviations: A/C, amoxicillin/clavulanate; T/S, trimethoprim/sulfamethoxazole; CZ, cefazolin; CIP, ciprofloxacin; NIT, nitrofurantoin; CTR, ceftriaxone).

band sizes (Supplementary Figure 1A and 1C) or signal (Supplementary Figure 1B), respectively, against the target DNA. We evaluated the ability of the SNP primers to recognize the corresponding SNPs in the background of 582 isolates with 60 *fimC* alleles and 156 *fimH* alleles that encompassed all naturally occurring combinations of base pairs that are variable within the primer-annealing regions (see Material and Methods). The presence or absence of the targeted SNP could not be correctly identified by PCR in only a few of the variable *fimC* or *fimH* alleles, and these error-causing alleles were very rare in the reference set of 2559 isolates (projected at 2.1%).

To confirm the specificity of designed primers towards *E. coli*, we tested them using isolates of various *Enterobacteriaceae* species, including (number of isolates) *Klebsiella pneumoniae* [10], *Klebsiella oxytoca* [4], *Citrobacter freundii* [5] and *Citrobacter koserii* [2], *Enterobacter aerogenes* [12], *Proteus mirabilis* [6], *Pseudomonas aeruginosa* [3], *Serratia marsenscens* [3], and *Morhanel-la morgani* [5]. Both the SNP-specific reactions and the *uidA* control produced negative results in all non-*E. coli* isolates, with the exception of 1 isolate designated by 16S typing as *E. aerogenes*. Thus, the newly designed SNP-specific primers demonstrate a robust ability to distinguish the targeted SNPs despite some background sequence variation in the primer-annealing regions.

The 7-Single Nucleotide Polymorphism Typing Test Can Identify *Escherichia coli* Clonotypes Directly From Urine Specimens

In addition to the validation using pure bacterial cultures, the qPCR-based 7-SNP typing test was performed on bacterial DNA obtained from 160 clinical urine samples, of which 77

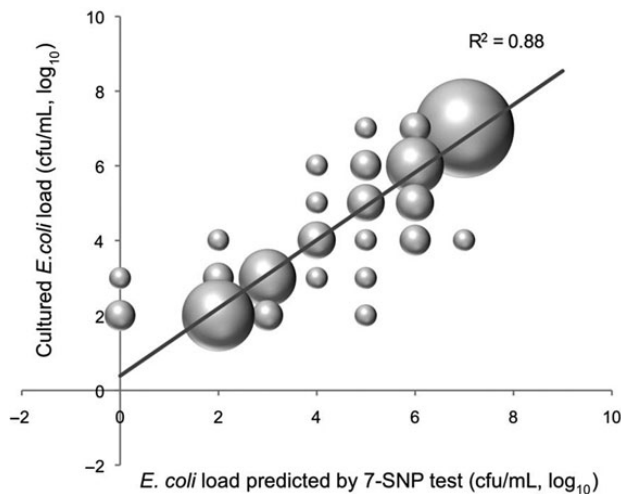


Figure 4. Detection of *Escherichia coli* in urine by 7-single nucleotide polymorphism (SNP) typing vs culturing. Total of 77 urine samples that had positive for *E. coli* growth are plotted here, with the *E. coli* load determined by 7-SNP test in quantitative polymerase chain reaction on the Y-axis and the culture-derived *E. coli* load on the X-axis. The size of bubbles is directly proportional to the number of urine samples with each combination of determined colony-forming units per milliliter (cfu/mL). The regression line represents the linear least square fit, with $\beta = 0.97$, $R^2 = 0.88$, and $P < .0001$.

were culture positive and 83 culture negative for *E. coli*. With the 83 *E. coli* culture-negative samples, the 7-SNP test was positive in 2 (98% specificity). With the 77 *E. coli* culture-positive urines, the 7-SNP typing test was positive in 74 (96% sensitivity), and the identified clonotypes matched the cultured isolate for all but 1 sample. Furthermore, in all 49 samples with $\geq 10^4$ colony-forming units (cfu)/mL *E. coli*, the 7-SNP test was positive and detected the same clonotype as was isolated from culture. Overall, the bacterial load determined by culture correlated closely with that predicted by the qPCR test based on the *E. coli*-specific *uidA* probe (Figure 4). In all 3 qPCR-negative but culture-positive urine samples, the bacterial load was low, ie, 10^2 cfu/mL (2 samples) or 10^3 cfu/mL (1 sample). Overall, multiple *E. coli* strains were encountered in $<1\%$ of specimens. Thus, the 7-SNP typing test can detect *E. coli* in urine samples and can reliably identify the corresponding clonotype, performing essentially as well quantitatively as the standard culture method.

DISCUSSION

Antimicrobial resistance is one of the most challenging problems in medicine today. Preculture, in-specimen identification of *E. coli* to the clonal group level could significantly improve empirical antibiotic selection [3]. For almost 2 decades, MLST has been the preferred method for determining the clonality of bacterial species using portable and easy-to-standardize sequence analysis [7]. However, MLST still requires sequencing 7 different housekeeping genes, making it relatively slow, costly,

and labor intensive [14]. We previously showed that clonal typing of *E. coli* can be done effectively and efficiently by sequencing just 2 genes, *fumC* and *fimH* [3]. *fumC* is part of the MLST scheme and is mostly under neutral selection. *fimH* encodes the type 1 fimbrial adhesin, which is present in $>95\%$ of all *E. coli* and undergoes diversifying selection for amino acid changes [3, 8]. The benefits of CH typing include the strong association of the resulting clonotypes with divergent antimicrobial susceptibility profiles, which recommends CH typing as a more rational basis for empirical antimicrobial therapy selection than whole-species antibiotic susceptibility patterns. However, because CH typing involves sequencing, and even the fastest pyrosequencing-based method requires at least 5 hours [3], CH typing is unlikely to be able to guide initial antimicrobial selection in the clinical setting.

The obvious approach for a fast, cost-effective, high-throughput typing method is detection of either specific gene regions or SNPs using DNA amplification. Multiplex tests were developed to determine *E. coli* phylogroups or STs based on amplification of specific genes [15, 16] or combinations of SNPs [17, 18]. Because CH typing provides greater diversity of *E. coli* clones and allows improved prediction of antibacterial susceptibility over MLST, we concentrated our search for clonotype-specific combinations of 7 SNPs selected from *fumC* and *fimH*.

By analysis of the *fumC* and *fimH* alleles present in 2559 isolates, we identified a combination of 7 SNPs that, based on computer analysis, separates the isolates into 54 clonotypes. Because extraintestinal *E. coli*, like most pathogens, are highly clonal in nature, the vast majority of strains belonged to a rather limited number of relatively large clonal groups. A high prevalence of relatively few dominant clonal groups in *E. coli* was shown by using MLST. However, due in part to the ability of 7-SNP typing to subdivide some of the largest STs into smaller groups, 7-SNP typing split the collection more evenly than did MLST. Moreover, this splitting of some STs was significant from a clinical diagnostics perspective. In particular, the largest ST, ST131, was split into 3 septatypes: 561, 560, and 510. Septatype 561 corresponded to the ST131-H30 subclone, which is almost exclusively responsible for the signature fluoroquinolone resistance of ST131 isolates and their CTX-M-15-mediated extended-spectrum β -lactamase ESBL phenotype. Another large ST, ST73, was split into septatype 620, corresponding to a high-virulence clonal group that includes model urosepsis isolate CFT073 [19], and septatype 661, corresponding to a clonal group that includes probiotic strain Nissle 1917 [20].

The most significant advantage of the 7-SNP typing test over typing methods requiring nucleotide sequence determination is its rapidity. Single nucleotide polymorphism-specific PCR primers recognize SNP presence or absence with a high degree of accuracy, using either conventional or real-time PCR, in a single- or multiplex configuration. The qPCR 7-SNP typing test with *uidA* control primers can identify *E. coli* within

20–25 minutes and is able to clearly distinguish *E coli* from other members of family *Enterobacteriaceae*.

Finally, we have demonstrated the power of the 7-SNP qPCR typing test for detection and clonotype determination of *E coli* directly from urine specimens. The entire procedure is technically simple and takes less than 1 hour to perform, and urine containing clinically significant numbers of *E coli* can be clonotyped within 30–40 minutes from the time of specimen collection. Test results correlate well with the results of standard culture, and the clonal type identified in urine match that of the cultured *E coli* isolate in 98% of samples overall, and in 100% of samples with clinically significant numbers of *E coli*. Rare instances of mismatch between clonotypes detected in the specimen versus from cultured *E coli* involved specimens containing low ($\geq 10^3$ cfu/mL) organism burdens, which approaches the 10^2 cfu/mL detection limit of both methods.

The 7-SNP test offers the potential to determine the presence of uropathogenic *E coli* septatypes prior to the availability of culture results and, thereby, to guide empirical antibiotic selection. In many locales, the prevalence of resistance to traditional mainstay antibiotics such as trimethoprim/sulfamethoxazole, amoxicillin/clavulanate, first-generation cephalosporins, and ciprofloxacin exceeds the Infectious Diseases Society of America-recommended 20% threshold for empirical use in treating uncomplicated cystitis in women [21]. In contrast, for the susceptible septatypes, the prevalence of resistance to each of these agents was well below this 20% cutoff. Thus, empirical antibiotic selection would be more rational if based on septatype rather than suspected species identity and species-based cumulative antibiograms.

CONCLUSIONS

In conclusion, the novel 7-SNP typing test described here should be a useful tool for both epidemiological analysis and clinical diagnostics, particularly for the prediction of antimicrobial susceptibility. It is highly reliable, inexpensive, easy to perform, robust, and can potentially be adapted for high-throughput automation. The test can be run in multiple instrumentation formats and performed on DNA obtained either from cultured bacteria or directly in clinical specimens.

Supplementary Material

Supplementary material is available online at Open Forum Infectious Diseases (<http://OpenForumInfectiousDiseases.oxfordjournals.org/>).

Acknowledgments

We thank the staff of clinical laboratories in the United States and abroad for their help in collecting the specimens and *E coli* isolates from patients. We also thank Helen Ghirmai Haile for technical assistance and Stephen Moseley and Dagmara Kisiela for critical evaluation of the manuscript.

Financial support. This material was supported by the National Institutes of Health (grant nos. R01AI106007 [to E. V. S.] and R41AI116114 [to

ID Genomics, Inc]) and the Office of Research and Development, Medical Research Service, Department of Veterans Affairs (grant no. 1 I01 CX000192 01; to J. R. J.).

Potential conflicts of interest. J. R. J. has received contracts, grants, or consultancies from Actavis, ICET, Janssen, Merck, and Tetraphase. J. R. J., E. V. S., S. C., P. A., and V. T. have patent applications pertaining to tests for specific *Escherichia coli* strains. E. V. S. is a founder and major shareholder in ID Genomics, Inc. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Foxman B. The epidemiology of urinary tract infection. *Nat Rev Urol* **2010**; 7:653–60.
2. Kaper JB, Nataro JP, Mobley HL. Pathogenic *Escherichia coli*. *Nat Rev Microbiol* **2004**; 2:123–40.
3. Tchesnokova V, Billig M, Chattopadhyay S, et al. Predictive diagnostics for *Escherichia coli* infections based on the clonal association of antimicrobial resistance and clinical outcome. *J Clin Microbiol* **2013**; 51:2991–9.
4. Wirth T, Falush D, Lan R, et al. Sex and virulence in *Escherichia coli*: an evolutionary perspective. *Mol Microbiol* **2006**; 60:1136–51.
5. Johnson JR, Murray AC, Kuskowski MA, et al. Distribution and characteristics of *Escherichia coli* clonal group A. *Emerg Infect Dis* **2005**; 11:141–5.
6. Tartof SY, Solberg OD, Manges AR, Riley LW. Analysis of a uropathogenic *Escherichia coli* clonal group by multilocus sequence typing. *J Clin Microbiol* **2005**; 43:5860–4.
7. Maiden MC, Bygraves JA, Feil E, et al. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci U S A* **1998**; 95:3140–5.
8. Weissman SJ, Johnson JR, Tchesnokova V, et al. High-resolution two-locus clonal typing of extraintestinal pathogenic *Escherichia coli*. *Appl Environ Microbiol* **2012**; 78:1353–60.
9. Feil EJ, Li BC, Aanensen DM, et al. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *J Bacteriol* **2004**; 186:1518–30.
10. Willcox WR, Lapage SP, Holmes B. A review of numerical methods in bacterial identification. *Antonie Van Leeuwenhoek* **1980**; 46:233–99.
11. CLSI. *Performance Standards for Antimicrobial Susceptibility Testing; Twentieth Informational Supplement*. Wayne, PA: Clinical and Laboratory Standards Institute; **2010**; 30:M100–S20.
12. Johnson JR, Tchesnokova V, Johnston B, et al. Abrupt emergence of a single dominant multidrug-resistant strain of *Escherichia coli*. *J Infect Dis* **2013**; 207:919–28.
13. Weissman SJ, Beskhlebnaya V, Chesnokova V, et al. Differential stability and trade-off effects of pathoadaptive mutations in the *Escherichia coli* FimH adhesin. *Infect Immun* **2007**; 75:3548–55.
14. Perez-Losada M, Cabezas P, Castro-Nallar E, Crandall KA. Pathogen typing in the genomics era: MLST and the future of molecular epidemiology. *Infect Genet Evol* **2013**; 16:38–53.
15. Clermont O, Christenson JK, Denamur E, Gordon DM. The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environ Microbiol Rep* **2013**; 5:58–65.
16. Clermont O, Bonacorsi S, Bingen E. Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Appl Environ Microbiol* **2000**; 66:4555–8.
17. Hommais F, Pereira S, Acquaviva C, et al. Single-nucleotide polymorphism phylotyping of *Escherichia coli*. *Appl Environ Microbiol* **2005**; 71:4784–92.
18. Sheludchenko MS, Huygens F, Hargreaves MH. Highly discriminatory single-nucleotide polymorphism interrogation of *Escherichia coli* by use of allele-specific real-time PCR and eBURST analysis. *Appl Environ Microbiol* **2010**; 76:4337–45.
19. Mobley HL, Green DM, Trifillis AL, et al. Pylonephritogenic *Escherichia coli* and killing of cultured human renal proximal tubular epithelial cells: role of hemolysin in some strains. *Infect Immun* **1990**; 58:1281–9.
20. Grozdanov L, Raasch C, Schulze J, et al. Analysis of the genome structure of the nonpathogenic probiotic *Escherichia coli* strain Nissle 1917. *J Bacteriol* **2004**; 186:5432–41.
21. IDSA. Principles and Strategies Intended to Limit the Impact of Antimicrobial Resistance, **2006**. Available at: http://www.idsociety.org/View_All_Statements_on_Antimicrobial_Stewardship/. Accessed 24 April 2006.