Journal of Infection (2015) 71, 615–626

BIA
British Infection Association

www.elsevierhealth.com/journals/jinf

REVIEW

Extra-intestinal pathogenic *Escherichia coli* (ExPEC): Disease, carriage and clones

Adam P. Dale [a,*], **Neil Woodford** [b,c]

[a] *Academic Unit of Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton General Hospital, South Academic Block, Tremona Road, Southampton SO16 6YD, UK*
[b] *Antimicrobial Resistance and Healthcare Associated Infections (AMRHAI) Reference Unit, Reference Microbiology Services, Public Health England, London NW9 5EQ, UK*
[c] *The NIHR Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance at Imperial College London, Hammersmith Campus, Du Cane Road, London W12 0HS, UK*

Accepted 16 September 2015
Available online 26 September 2015

KEYWORDS
Escherichia coli;
ExPEC;
Virulence factors;
Clones

Summary Extra-intestinal pathogenic *Escherichia coli* (ExPEC) have a complex phylogeny, broad virulence factor (VF) armament and significant genomic plasticity, and are associated with a spectrum of host infective syndromes ranging from simple urinary tract infection to life-threatening bacteraemia. Their importance as pathogens has come to the fore in recent years, particularly in the context of the global emergence of hyper-virulent and antibiotic resistant strains. Despite this, the mechanisms underlying ExPEC transmission dynamics and clonal selection remain poorly understood. Large-scale epidemiological and clinical studies are urgently required to ascertain the mechanisms underlying these processes to enable the development of novel evidence-based preventative and therapeutic strategies. In the current review, we provide a concise summary of the methods utilised for ExPEC phylogenetic delineation before exploring in detail the associations between ExPEC VFs and site-specific disease. We then consider the role of ExPEC as an intestinal colonist and outline known associations between ExPEC clonal variation, specific disease syndromes and antibiotic resistance.

Introduction

Escherichia coli is a genetically diverse species comprising non-pathogenic gut commensals and strains responsible for intestinal and extra-intestinal disease. While awareness amongst healthcare professionals and the public relating to *E. coli* strains associated with intestinal disease, *e.g.* enterohaemorrhagic *E. coli* (serogroup O157:H7), is high, the same

* Corresponding author. Tel.: +44 (0)23 81206455.
E-mail address: A.P.Dale@Soton.ac.uk (A.P. Dale).

http://dx.doi.org/10.1016/j.jinf.2015.09.009

has not always been true for strains associated with extra-intestinal disease.[1] In the context of rapidly increasing multidrug-resistance worldwide and a diminishingly effective antimicrobial arsenal to tackle resistant strains, extra-intestinal *E. coli* infections are now a serious international public health concern associated with a significant economic impact. Consequently, there is an increasing awareness of the importance of extra-intestinal *E. coli* amongst both healthcare professionals and the general public alike.[2,3]

Historically, extra-intestinal *E. coli* isolates were separated into groups determined by disease association, including uropathogenic *E. coli* (UPEC), neonatal meningitis-associated *E. coli* (NMEC), and sepsis-causing *E. coli* (SEPEC). These terms were subsequently deemed outdated following the observation that *E. coli* isolates assigned to specific groups were capable of causing infection at multiple anatomical sites. In 2000, Russo and Johnson suggested the term extra-intestinal pathogenic *E. coli* (ExPEC) as an alternative descriptor for all non-commensal *E. coli* isolates capable of causing extra-intestinal disease.[4] Unlike commensal *E. coli*, ExPEC have the ability to cause disease once outside the host gut reservoir due to the possession of pathogenic virulence factors. From a molecular viewpoint, Peirano et al.[5] defined ExPEC as isolates containing at least two of the following virulence factors (VFs) within their genome: *papA* and/or *papC*, *sfa/foc*, *afa/draBC*, *kpsM II* and *iutA* (see Table 1).

ExPEC are most frequently implicated as urinary pathogens and are isolated as the infective agent in up to 90% of both simple community-acquired urinary tract infections (UTIs) and pyelonephritis cases. Other infections of the urinary tract, including prostatitis and catheter-associated UTIs, are also frequently caused by ExPEC.[6–8] ExPEC are frequently implicated in infections originating from abdominal and pelvic sources including, but not limited to, biliary infections, infective peritonitis, and pelvic inflammatory disease.[9,10] Less frequently, they are associated with skin and soft tissue infections, neonatal meningitis and hospital-acquired pneumonia.[11,12] Haematogenous invasion of ExPEC from the initial infective focus results in the sepsis syndrome which, in the absence of timely management, may result in death.

Following a brief overview of the methods used for ExPEC phylogenetic delineation, the current review will consider the role of ExPEC as a potential intestinal commensal and detail the associations between ExPEC strains, their virulence factor profiles, and host disease.

ExPEC lineage determination; *a brief history*

Methodologies used to define and understand ExPEC lineages and *E. coli* phylogeny are numerous and have evolved in parallel with the availability of new technologies (see Table 2). In-depth analyses are available elsewhere.[13,14] Briefly, in the early 1970s, O antigen-based serotyping, followed later by the addition of H and K antigen serotyping, was first utilised to delineate *E. coli* isolates from humans and other animals and allowed the identification of some of the *E. coli* strains we now refer to as ExPEC.[15,16] In 1984, the pioneering work of Ochman and Selander led to the establishment of the *E. coli* reference strain collection (ECOR) comprising 72 isolates from human and other mammalian hosts.[17] Multi-locus enzyme electrophoresis (MLEE) separated the isolates into five key phylogenetic groups (phylogroups), namely A, B1, B2, D and E. The distribution of ExPEC isolates within these phylogroups will be discussed in detail later, however they are mainly limited to groups B2 and D (see Fig. 1).

In 2000, a triplex polymerase chain reaction (PCR)-based method was devised by Clermont et al. and enabled rapid *E. coli* phylogroup assignment.[18] This methodology was much faster than MLEE, was simple and inexpensive, and allowed *E. coli* isolates to be separated into four main phylogroups (A, B1, B2, D). An updated multiplex PCR method was more recently devised by Doumith et al.[19] An independent study by Turrientes et al. subsequently compared both methods against a multi-locus sequence typing standard and demonstrated superiority of Doumith's method with regards to accuracy of phylogroup assignment.[20]

Since 2000, the understanding of *E. coli* phylogeny has improved significantly. Eight *E. coli* phylogroups are currently recognise d (A, B1, B2, D, E, F, G, and Clade 1) and a new PCR approach enables isolates to be assigned to one of these phylogroups.[21]

As DNA sequencing methods became more widely available, they superseded MLEE as the preferred technologies for phylogenetic analysis given their superior discriminative ability. Multi-locus sequence typing (MLST) involves sequencing of selected (often seven) bacterial housekeeping genes and, due to its standardised approach and greater resolution as compared with 'phylogrouping', has allowed more detailed analysis of ExPEC lineages.[22,23] It separates the isolates into distinct sequence types (STs), which are defined as isolates with identical allelic profiles, and into broader clonal complexes (CCs), which are defined as a group of at least three STs each differing from the others by no more than 1 of 7 alleles.[24] Of the three MLST schemes available for *E. coli*, the Achtman scheme (http://mlst.warwick.ac.uk/mlst/dbs/Ecoli) is most commonly used.[14] *E. coli* sequence type (ST) data presented in the current review are derived from the Achtman scheme unless otherwise stated.

Although MLST is the preferred method for determining phylogenetic relationships in ExPEC, the discriminatory power of this technique is limited. Isolates belonging to the same ST can be genetically distinct and may be associated with variable pathotypic behaviours. In 2012, Weissman et al.[25] described a new method, CH typing, which derives its name from *fumC* and *fimH* gene analysis. They demonstrated that this approach was able not only to predict the respective MLST-based profile with up to 95% accuracy, but that it also enabled large STs to be split into a number of smaller clonal subgroups.[25] Although CH typing will not replace MLST as a tool for phylogenetic studies, there are clear advantages of this technique, particularly relating to delineation of clones within STs and reduced costs when performing preliminary evaluations on larger clinical specimen collections.[25,26]

Ultimately, techniques such as pulsed-field gel electrophoresis (PFGE) and whole genome sequencing (WGS) provide the greatest resolution for the purposes of ExPEC phylogenetic analysis, such as may be needed for outbreak