Abstract

The 2011 *Escherichia coli* outbreak in Germany, which resulted in more than 4000 cases including 908 cases of hemolytic-uremic syndrome (HUS) and at least 50 deaths, highlighted the genome plasticity of *E. coli* and the potential for new virulent strains to emerge. The analysis of 170 *E. coli* genome sequences for the presence of nine previously identified protective extraintestinal pathogenic *E. coli* antigens, suggested the feasibility of a combination vaccine as a universal intervention against all pathogenic *E. coli*. 
*Escherichia coli* is part of the human and animal commensal flora, lives in several different environments but is also a well-known pathogen. In 2011, an outbreak associated with a newly emerged pathogenic *E. coli* originated in Germany and caused 908 cases of hemolytic-uremic syndrome (HUS) and 3,167 non-HUS cases in Europe, leading to 50 deaths (1). Based on the site of infection and disease caused by *E. coli* in humans, pathogenic *E. coli* strains are divided in two major groups: extraintestinal pathogenic *E. coli* (ExPEC) and intestinal pathogenic *E. coli* (InPEC). ExPEC strains are responsible for disease outside the intestinal tract such as urinary tract infections (UTIs), sepsis and meningitis, and are classified as uropathogenic *E. coli* (UPEC) or neonatal meningitis-associated *E. coli* (NMEC). Among the InPEC strains causing diarrhoeagenic infections, several well-defined pathotypes have been identified including: enteropathogenic *E. coli* (EPEC); enterotoxigenic *E. coli* (ETEC); enterohaemorrhagic *E. coli* (EHEC); enteroaggregative *E. coli* (EAEC); enteroinvasive *E. coli* (EIEC); and adherent-invasive *E. coli* (AIEC).

The strains responsible for the 2011 German outbreak belong to the O104:H4 serotype and caused disease typical of EHEC infections, which are characterized by destruction of gut epithelium cells due to the release of Shiga toxin (Stx), resulting in abdominal cramps, bloody diarrhea and the triggering of HUS. However, genome sequencing of the outbreak strains (LB226692, TY2482 and H112180280) and Multi Locus Sequence Typing (MLST) analysis revealed that they are closely related to the EAEC strain 55989, as shown by the localization of these strains on the same branch in Figure 1. EAEC strains use aggregative adherence fimbriae (AAF) to stimulate a strong interleukin-8 response and biofilm formation on the surface of epithelial cells, followed by disruption of actin cytoskeleton and exfoliation by a serine protease autotransporter (Pet) (2).

The 2011 outbreak strains perfectly illustrate the genome plasticity and evolution of *E. coli* as a result of horizontal gene transfer. These strains combine the virulence mechanisms of two pathotypes (EAEC and EHEC), leading to an improved ability to adhere to and infect host cells. Furthermore, the acquisition of mechanisms mediating increased antibiotic resistance hampered patient treatment and recovery. These strains have conserved most of the virulence factors of an EAEC strain, but several mobile genetic elements (bacteriophages, plasmids, transposons and pathogenicity islands) were responsible for the acquisition of new attributes, including the phage-mediated Stx2a, extended-spectrum beta-lactamases, tellurite and mercury resistance genes, type IV pilus system, colicin, hemolysin E, SPATE proteases, and F fimbriae, a system involved in high frequency recombination, mobilization and transfer of genes (3, 4). In 2001, a similar O104:H4 strain, 01-09591, combining the features of EHEC and EAEC pathotypes and expressing both AAF and Stx was isolated in Germany from patients with HUS (5).

For many years, *E. coli* pathotypes have been studied in isolation instead of addressing *E. coli* as a single microorganism responsible for human and animal diseases. As seen by analysis of
the most recent outbreak strains, *E. coli* uses several mechanisms to exchange genetic material and novel strains representing combinations of different pathotypes may emerge in the future. This variability and adaptability reinforces the need for a novel approach to combat pathogenic *E. coli*. Considering the increasing antibiotic resistance present among *E. coli* strains, which is derived from an uncontrolled use of antibiotics in humans and in the veterinary field, vaccination is the most promising approach to control disease. Comparative genome analysis and reverse vaccinology led to the identification of nine antigens capable of inducing protective immune responses against ExPEC strains, several of which are also prevalent in different intestinal *E. coli* pathotypes (6). The feasibility of a universal intervention against all pathogenic *E. coli* is highlighted in Figure 1, which shows that all *E. coli* strains sequenced to date could be covered by a vaccine that contains at least four antigens: ECOK1_0290, ECOK1_3385, ECOK1_3457 and c0975. The first two of these antigens have been further characterized, and ECOK1_0290 is a broadly conserved adhesin, renamed FdeC (Factor adherence *E. coli*), that elicits protection in mouse sepsis and mouse urinary tract infection models (6, 7). ECOK1_3385, a putative metalloprotease, is surface-exposed and secreted by a novel type two secretion system (T2SS), that is able to confer nearly complete protection from bacteraemia and mortality in a murine model of sepsis after either active or passive immunization (6). ECOK1_3457 is involved in iron acquisition (8), and c0975 is annotated as a hypothetical protein. The ability of these antigens to confer protection also against intestinal pathotypes, the route of administration and the duration of the immune response will need further investigation. In addition, the effect of vaccinating with antigens, which in some cases are also present in *E. coli* commensals, on the composition of the natural intestinal flora requires additional evaluation. However Proteobacteria (including *E. coli*) represent less than 0.1% of the human flora (9).

In conclusion, we propose that *E. coli* should be treated as a single microorganism capable of causing varied diseases in both humans and animals. Despite the alternative mechanisms that have evolved to colonize and adapt to new niches, *E. coli* strains have maintained a core genome sequence and therefore share several components that could be useful targets for a universal vaccine against *E. coli*. From an evolutionary point of view, any commensal or environmental isolate has the potential to acquire novel virulence factors and become a pathogenic strain, and the continuous exchange of genetic material between pathotypes could impact the future coverage and efficacy of a vaccine against *E. coli*. Therefore we need to consider *E. coli* as a microorganism that is continuously evolving, and look for highly represented antigens that, in combination, could provide an effective vaccine that would prevent outbreaks occurring in the future.

**Conflict of interest/ Statement of financial support:** The authors are employees of Novartis Vaccines and Diagnostics.
References


Figure 1 – The evolutionary relationship and distribution of protective antigens among sequenced *E. coli* strains. The evolutionary history of 170 strains was inferred from MLST data using the Neighbor-Joining method. The presence of nine protective antigens, identified from Reverse Vaccinology of ExPEC strains, are shown as colored squares and are sorted from the most represented (inner circle) to the least represented (outer circle). Strains were tentatively classified when possible as InPEC (EHEC, ETEC, EIEC, EAEC, AIEC, EPEC colored as shown in the legend), ExPEC (NMEC, UPEC, APEC) and mainly fecal or environmental isolates (not colored).