

**Functional genotyping of extraintestinal pathogenic *E.coli* (ExPEC)
belonging to the highly pathogenic Sequence type 95 reveals the zoonotic
nature of human and avian strains**

Inaugural-Dissertation

To obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

Submitted to the Department of Biology, Chemistry and Pharmacy
Of Freie University Berlin

By

Nishant Nandanwar

From (Bhopal, M.P., India)

2013

Time Period – November 2010 till October 2013

Supervisor – Prof. Dr. Lothar H. Wieler

Institute – Institute for Microbiology and Epizootics,

Robert-von-Ostertag-Str. 7-13,

14163, Berlin, Germany.

1st Reviewer: Prof. Dr. Lothar H. Wieler

2nd Reviewer: Prof. Dr. Heribert Hofer

Date of Defense: 8th April 2014

Acknowledgement

Above all, I wish to thank God and my parents for their abundant blessings and graces throughout my life and during my doctoral work.

I would like to express my sincere thanks to Prof. Dr. Lothar Wieler for giving me the opportunity to pursue my doctorate under his supervision. I am very thankful to him for his constant support and guidance during this work, and for his constant encouragement and motivation.

I would also like to thank Prof. Dr. Christa Ewers in a special way for her supervision, guidance and support right from the beginning of my work.

My special thanks to Prof. Dr. Heribert Hofer, Leibniz Institute for Zoo and Wildlife Research (IZW), Berlin, for his external supervision during my doctoral studies.

I would like to express my sincere gratitude to Associate Prof. Dr. Niyaz Ahmed, Department of Biotechnology & Bioinformatics, University of Hyderabad, India for his immense support and for providing me the opportunity to pursue my doctorate in Germany.

Special thanks to Prof. Dr. Ulrich Dobrindt, Institut für Hygiene, Zentrum für Klinisch-Theoretische Medizin, Münster, for providing me with UTI89 strain.

I would like to thank Torsten Semmler, Dr. Haritha Devi for helping me with the bioinformatic analyses during my thesis work.

I would also like to express my gratitude to Dr. Traute Janßen, for her support to organize and carryout my experiments successfully.

I would like to thank Dr. Esther-Maria Antão for her support, motivation and taking care of credit points and updates for interesting courses during this work.

I am thankful to Dr. Karsten Tedin for correcting my thesis.

Thanks to Sarah Brüggemann for her constant support during my work.

I am thankful to Inga Eichorn and Szilvia Vincze for their support and helping me with the translations from English to German.

Thanks to, Stefan Herbel Vikash Singh, Carmen Bednorz for their help and support.

My special thanks to Monika Feldhan, Michael Kühl, Kathrin Oelgeschläger, Petra Krienke, Bianca Kinnemann and other colleagues at the institute for their help and friendliness all the time.

I would also like to thank my colleagues Ashutosh Kumar, Savita Devi, Arif Hussain, Kishore Nalam from India as their regular visit did not allow me to sense loneliness and for their motivation during my thesis work.

Aim of this project

Escherichia coli represent a bacterial species of high relevance for human and animal health. A subgroup of *E. coli* are Extraintestinal pathogenic *E. coli* (ExPEC), a definition which is based on the site of infection or the habitat, namely outside of the intestine, and the expression of certain virulence-associated genes. ExPEC frequently cause urinary tract infection (UTI), septicemia and meningitis in human and animals. In recent years great efforts have been made to discern the group of ExPEC with respect to clinical implications and host origin, but still a precise definition and thus a sound risk assessment of these strains, in particular with regard to their zoonotic potential, is not found.

Sequencing based methods like single locus sequence typing (SLST) and multilocus sequence typing (MLST) are frequently used to study the epidemiology, infectious biology and transmission pathways of bacterial pathogens (Singh et al. 2006, Maiden et al. 1998, Wirth et al. 2006). MLST compares the nucleotide sequences of internal 400- to 500-bp regions of seven housekeeping genes supposed to be under stabilizing selection (Tartof et al. 2007). Each bacterial isolate is defined by distinct allele combinations of these genes and based on that a sequence type (ST) assigned. Isolates sharing their allelic profile belong to the same ST/phylogroup, while strains differing in no more than one allele to their direct neighbor form ST complexes (STCs). A previous screen of 1,030 ExPEC strains from various hosts and clinical sources via MLST by our group revealed that these pathogens present a highly diverse group of strains, while some STs appeared over represented in the population (Fig. 1). One prominent phylogroup is ST95, which accumulates highly virulent strains having a wide spectrum of virulence-associated genes and comprising strains mainly of human and avian origin. Although *E. coli* strains from dogs and cats represented a proportion of nearly 25% of the entire ExPEC collection, ST95 strains could only be isolated from two dogs. In addition, one strain from an urban rat was identified belonging to ST95. Thus, ST95 resembles a group of strains which may be termed as “Human-Avian” complex, based on previous results from our group, where a large set of strains from cattle, swine, dogs and cats were mostly determined as non-ST95 strains, substantiating the limited host range of ST95 strains (Ewers et al unpublished data). As according to the publicly available database (<http://mlst.ucc.ie/mlst/mlst/dbs/Ecoli/> as observed on April 2013) ST95 currently contains strains, mostly of human and avian origin with an exception of one strain isolated from dog,

providing support for the existence of some kind of host limitation of these strains. This is further evidenced, as most of the phylotypes (>2,000) identified so far are mixed groups of strains from different sources, including human, poultry and other livestock animals as well companion animals (dogs, cats, horses).

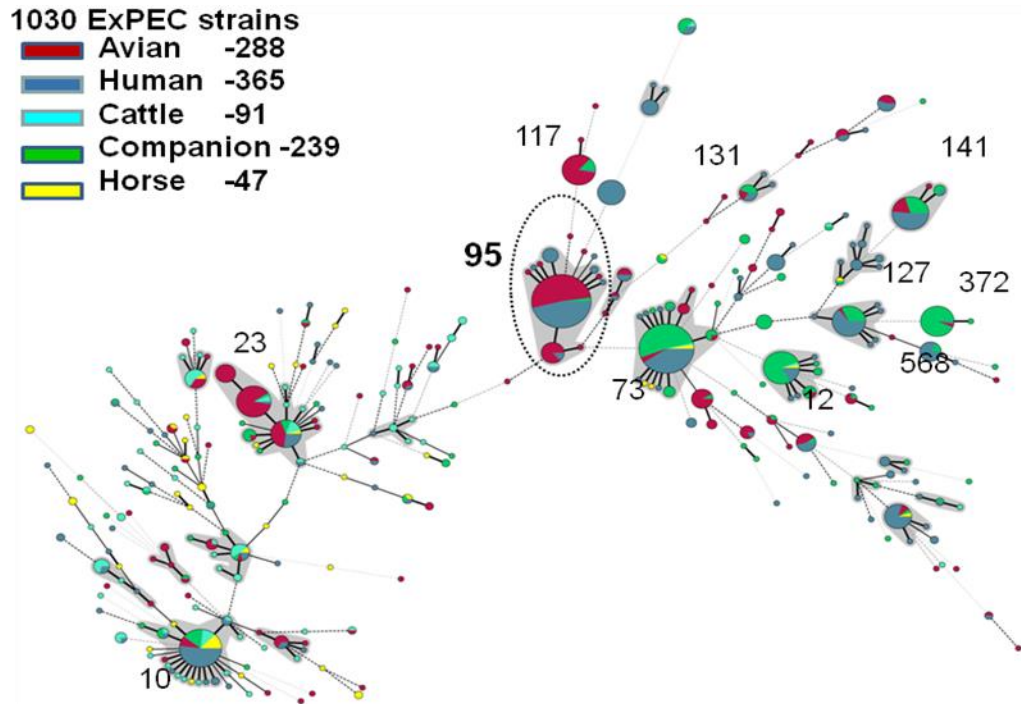


Figure 1: Minimum spanning Tree (MSTree) based on the allele combination of seven house-keeping genes of 1030 ExPEC strains (Ewers, Wieler et al., unpublished data). Each circle resembles a unique ST; the thickness of the circles represents the number of isolates included.

Thus, with this background following aims were setup for this project.

- I. Identification of genetic markers, so called single nucleotide polymorphism (SNP) that discern ST95 strains in a host specific manner by analyzing the core genome and flexible or accessory gene pool, thereby facilitating a sound risk assessment of this group of strains in future.
- II. Identification of certain phenotypic characteristics in strains from ST95 complex that would segregate this group of strains according to their specific human/avian host respectively.

1. Introduction	6
1.1 Commensal <i>E. coli</i>	6
1.2 Intestinal pathogenic <i>E. coli</i> (InPEC)	7
1.3 Extraintestinal pathogenic <i>E. coli</i> (ExPEC)	7
2. Virulence associated factors carried by ExPEC	9
2.1. Adhesins	9
2.1.1. Adhesins expressed by ExPEC	10
2.1.1.1. Type 1 fimbriae	11
2.1.1.2. P fimbriae	12
2.1.1.3. F1C fimbriae	13
2.1.1.4. S fimbriae	14
2.1.1.5. Curli fimbriae	14
2.1.2. Afimbrial adhesins expressed by ExPEC - Dr/Ara fimbriae	15
2.2. Polysaccharide coatings	16
2.3. Lipopolysaccharide (LPS)	16
2.4. Toxins and Proteases	17
2.4.1. Haemolysin	17
2.4.2. Cytotoxic necrotizing factor (CNF toxin)	18
2.4.3. SPATE (serine protease autotransporters of <i>Enterobacteriaceae</i>) family	19
2.5. Iron acquisition system	20
2.6. Serum resistance	22
3. Diseases related to ExPEC	24
3.1. Urinary tract infections	24
3.2. Neonatal meningitis	26
3.3. Sepsis	27
3.4. Avian systemic <i>E. coli</i> infection	28
4. Zoonotic nature of ExPEC	29
5. Multilocus sequence typing (MLST)	30
6. ExPEC related to sequence type complex (STC95)	31
2. Materials	33
2.1. Chemicals	33
2.2. Materials for microbiological experiments	33
2.2.a. Bacterial strains	33
2.2.b. Growth media	34
2.3. Materials for molecular biology experiments	35
2.4. Materials for cell biology experiments	39
2.5. Genomes for bioinformatic analysis	40
2.6. Statistical analyses	41
3. Methods	42
3.1. Microbiology experiments	42
3.1.a. Biofilm formation assay	42
3.1.a.i. 96 microtiter plate method	42
3.1.a.ii. Long term biofilm formation	43
3.1.b. Haemagglutination assay	43
3.1.c. Serum assay	43

3.2.	Molecular biology experiments	44
3.2.a.	Screening of available strains from different hosts for sequence type 95.....	44
3.2.b.	<i>fimH</i> gene amplification.....	45
3.2.c.	Cloning of prominent <i>fimH</i> alleles.....	45
3.2.c.i.	Extraction of Genomic DNA.....	46
3.2.c.ii.	Preparation of competent cells.....	46
3.2.c.iii.	Long range PCR for amplification of DNA insert.....	47
3.2.c.iv.	Ethanol precipitation of DNA.....	48
3.2.c.v.	Isolation of plasmid by alkaline lysis method.....	48
3.2.c.vi.	Restriction digestion of plasmid vector - PET 45b (+) and DNA insert.....	48
3.2.c.vii.	Purification of digested DNA and plasmid vector.....	49
3.2.c.viii.	Ligation of plasmid vector and DNA insert.....	49
3.2.c.ix.	Transformation of competent cells.....	49
3.2.c.x.	Screening of colonies for DNA inserts.....	50
3.2.c.xi.	Agarose gel electrophoresis.....	51
3.2.d.	TOPO cloning	51
3.3.	Cell culture experiments	52
3.3.a.	Adhesion and invasion assay	52
3.3.a.i)	Reviving of cell line from liquid nitrogen.....	52
3.3.a.ii)	Subculturing of adherent cells.....	52
3.3.a.iii)	Seeding the cells.....	53
3.3.a.iv)	Performing adhesion and invasion assay.....	53
3.4.	Bioinformatics analysis	54
3.4.a.	Comparison of maximum common genome (MCG) sequences of 46 different <i>E. coli</i> strains	54
3.4.b.	<i>In silico</i> analysis of 46 genomes of <i>E. coli</i> strains including STC95 strains, based on recombination and mutation events to identify candidate genes for SNP analysis discerning ST95 strains according to specific human/avian host	55
3.4.c.	Identification of genetic markers by analyzing genes of core genome from nine fully sequenced ST95 strains genome	56
3.4.d.	SNP analysis of <i>E. coli</i> fimbrial adhesin gene	58
4.	Results	59
4.1.	Screening of available strains from different animal hosts for sequence type 95	60
4.2.	Identification of genetic markers -- single nucleotide polymorphisms (SNPs) - that discern ST95 strains with respect to their host origin	61
4.2.a.	Comparison of maximum common genome (MCG) sequences of 46 different <i>E. coli</i> strains	61
4.2.b.	<i>In silico</i> analysis of 46 genomes of <i>E. coli</i> strains including STC95 strains, based on recombination and mutation events to identify candidate genes for SNP analysis discerning ST95 strains according to specific human/avian host origin ..	63
4.2.c.	Identification of genetic markers by analyzing the MCG of nine genomes of strains of ST95	64
4.2.d.	SNP analysis of <i>E. coli</i> fimbrial adhesin gene	67
4.3.	Cloning of prominent <i>fimH</i> alleles	69
4.4.	Identification of phenotypic characteristics in ST95 complex strains that could segregate this group of strains according to their human/avian host respectively	69

4.4.a Adhesion and invasion assay	69
4.4.b. Biofilm formation assay	74
4.4.b.i. 96 microtiter plate method.....	74
4.4.b.ii. Long term biofilm formation.....	76
4.4.c. Haemagglutination assay	80
4.4.d. Serum assay (Chicken and human serum)	81
5. Discussion	84
5.1. Screening of available strains from different animal hosts for sequence type 95	85
5.2. Identification of genetic markers, single nucleotide polymorphisms (SNPs) that discern ST95 strains with respect to their host origin	86
5.3. Identification of certain phenotypic characteristics in ST95 strains that could segregate this group of strains according to their human/avian host respectively	93
5.3.a. Adhesion and invasion capabilities of human and avian strains of ST95.....	95
5.3.b. Human and avian strains of ST95 represent similar biofilm forming capabilities...	96
5.3.c. Curli and cellulose expression (Long term biofilm formation).....	97
5.3.d. Haemagglutination assay among ST95 strains.....	99
5.3.e. Serum resistance of ST95 strains in chicken and human serum.....	100
6. Conclusion	102
7. Zusammenfassung	103
8. References	105

1. Introduction

Escherichia coli, a Gram negative bacterium, is normal inhabitant of the gut microbiota representing a bacterial species of high relevance for human and animal health and is associated with a diverse spectrum of diseases (Wang et al. 2009, Bauchart et al. 2010). *E. coli* is coexisting with its hosts all lifelong. It colonizes the gastrointestinal tract of human infants within few hours after birth (Kaper et al. 2004) and is acquired either from the environment or from the mother during parturition (Bokranz et al. 2005). *E. coli* is a diverse organism and certain pathovars of pathogenic *E. coli* cause a wide range of diseases affecting humans and animals worldwide (Croxen and Finlay 2010). Different *E. coli* pathovars adapt themselves by acquiring specific virulence factors conferring their ability to carve new niches, allowing them to cause wide spectrum of disease to their host (Kaper et al. 2004). According to the genetic and clinical perspective, *E. coli* can be grouped as (a) commensal *E. coli* (i. e. harmless intestinal dwellers), (b) intestinal pathogenic *E. coli* (i.e. enteric and diarrheagenic strains) and (c) extraintestinal pathogenic *E. coli*. Those strains that are responsible for extraintestinal infections are termed as extraintestinal pathogenic *E. coli* (ExPEC) (Russo and Johnson 2000). ExPEC frequently cause urinary tract infection (UTI), septicemia, meningitis and systemic infections in birds (Kariyawasam et al. 2006, Russo and Johnson 2003). ExPEC include uropathogenic *E. coli* (UPEC) associated with urinary tract infection in human and animals, neonatal meningitis *E. coli* (NMEC) as well as septicemic *E. coli* (SePEC) causing systemic infection in human and animals and avian pathogenic *E. coli* (APEC) that cause systemic infection in birds leading to economic losses in poultry industry (Babai et al. 1997, Ewers et al. 2004). *E. coli* have been phylogenetically grouped into six groups designated as A, B₁, B₂, C, D and E, and it has been shown that particularly D and B₂ harbor ExPEC (Smith et al. 2007).

1.1 Commensal *E. coli*

Commensal *E. coli* strains are members of the gastro intestinal microbiota of most mammalian host, including humans (Blyton et al. 2013). They are considered to play a role in supporting digestion and to provide defense mechanisms against enteric pathogens by successfully competing with other microbiota and hindering colonization of pathogenic agents and also produce vitamin K (Schierack et al. 2009) that is beneficial for the host. Commensal *E. coli* are rarely associated with disease except in immune compromised hosts and make their niche in the mucous layer of the mammalian colon (Kaper et al. 2004). Commensal strains of human origin

are mainly derived from phylogenetic group A and B1, typically lacking specialized virulence attributes that are found in intestinal pathogenic and extraintestinal pathogenic *E. coli* (Picard et al. 1999).

1.2 Intestinal Pathogenic *E. coli* (InPEC)

Intestinal pathogenic *Escherichia coli* (InPEC) are responsible for health problems in mammals, including humans (Muller et al. 2007). Based on the virulence factors, severity of clinical implications and prognosis, presently intestinal or diarrheagenic *E. coli* bacteria are grouped into six major pathotypes enteropathogenic *E. coli* (EPEC), enterohaemorrhagic *E. coli* (EHEC), enterotoxigenic *E. coli* (ETEC), enteroaggregative *E. coli* (EAEC), enteroinvasive *E. coli*, (EIEC) and diffusely adherent *E. coli* (DAEC) (Nataro and Kaper 1998). Strains from these pathotypes show diversity in phylogenetic groupings but are associated mainly with the A, B1, or D phylogenetic groups (Smith et al. 2007).

1.3 Extraintestinal Pathogenic *E.coli* (ExPEC)

ExPEC are those *E. coli* strains that cause extraintestinal infections and are responsible for a wide spectrum of diseases like urinary tract infections (UTI), newborn meningitis (NBM), sepsis and septicemia (Ron 2006). Additionally, an animal pathotype of extraintestinal pathogenic *E. coli* is avian pathogenic *E. coli* (APEC) causing respiratory infection, septicemia in poultry and together with other ExPEC strains account for huge health and economic lost (Kaper et al. 2004, Smith et al. 2007). Different categories of intestinal and extraintestinal pathogenic *E. coli* are depicted in Fig. 2.

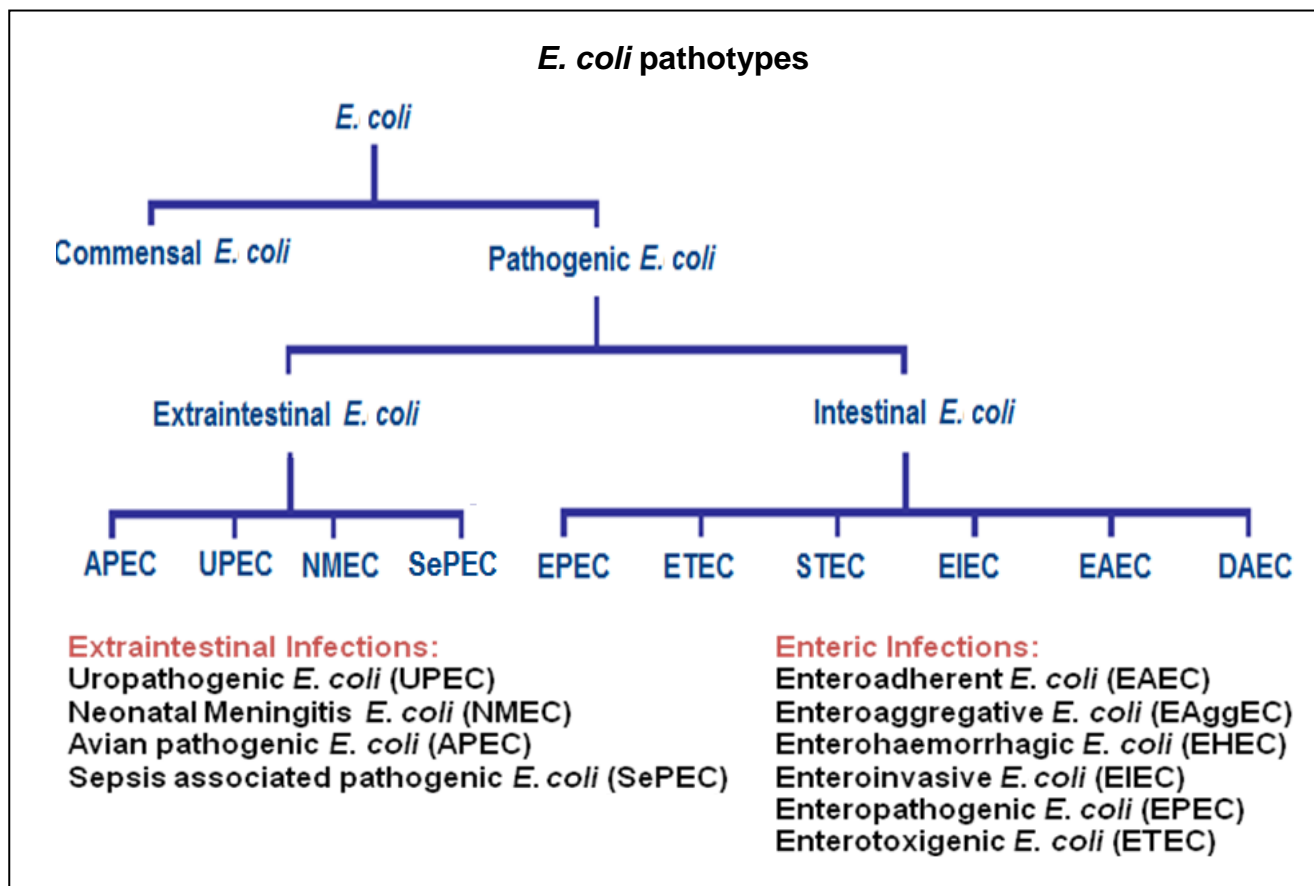


Figure 2: Schematic representation of *Escherichia coli* groups and pathotypes.

Phylogenetically and epidemiologically ExPEC are potentially different from those of intestinal pathogenic and commensal strains (Smith et al. 2007). Most of the ExPEC strains phylogenetically belong to B₂ and to a lesser extent D groups and are equipped with various virulence factors that help these strains during different mode of infection mechanisms like adhesion, invasion of host tissues, escape host defence mechanisms, signaling and production of different toxins interfering host cellular functions thereby promoting extraintestinal infection in both normal and immune compromised hosts (Dobrindt and Hacker 2008, Wiles et al. 2008). ExPEC clonal groups were previously identified partially by serotype O:K:H serotype (e.g. O18:K1:H7) (Orskov et al. 1982). ExPEC clonal groups are mainly derived from B2 phylogenetic group and to some extent group D explaining the prevalence of B2 and D phylogenetic group among ExPEC clinical isolates (Picard et al. 1993, Bingen et al. 1998).

ExPEC possess diverse virulence factors such as adhesins, invasins, polysaccharide coatings (e.g. lipopolysaccharides and capsules), toxins, siderophores, proteases and serum resistance

proteins. Many virulence factors are encoded on PAIs (Pathogenicity associated islands) that are blocks of contiguous virulence genes that are normally absent in commensal *E. coli* such as *E. coli* K-12 (Groisman and Ochman 1996, Johnson and Russo 2002, Ewers et al. 2007).

In 2003 Johnson et al. defined ExPEC strains as *E. coli* isolates containing two or more of the following virulence markers as determined by multiplex PCR: *papA* (P fimbriae structural subunit) and/or *papC* (P fimbriae assembly), *sfa/foc* (S and F1C fimbriae subunits), *afa/dra* (Dr-antigen-binding adhesins), *kpsMTII*(group 2 capsular polysaccharide units), and *iutA* (aerobactin receptor) (Johnson et al. 2003).

2. Virulence-associated factors carried by ExPEC

There are several virulence-associated factors (VAFs) that contribute to the pathogenicity of organisms by providing survival advantage to cope up with the unfavorable conditions within host and cause infection. Virulence factors can be grouped based on their functionality e.g. adhesins, invasins, siderophores systems, toxins, surface polysaccharides, serum resistance associated traits (Johnson and Russo 2005). Many VAFs occur together suggesting that there is a direct genetic linkage or co-selection (Johnson and Stell 2000, Kanamaru et al. 2003). It has been demonstrated that there is a genetic linkage of VAFs within pathogenicity islands (PAIs) and on plasmids (Guyer et al. 1998, Dobrindt et al. 2001).

2.1 Adhesins

Adherence is an important aspect that mediates colonization of the host tissue by pathogen. Colonization is mediated by the attachment of the bacterium to the receptors expressed by cells forming the lining of the mucosa (Wizemann et al. 1999). Successful colonization later helps microbial pathogens to invade host mucosal surface and reach tissue cells, thereby causing infection. Attachment to the host tissue helps pathogens to avoid from, being swept away by normal body fluid like urine in case of UTI infections and getting eliminated from the host (Johnson 1991). This attachment is mediated by adhesins that is associated with fimbriae and directs high affinity binding to specific cell surface components. Fimbriae are long hair like extracellular appendages that mediate specific attachment to the host epithelial cell surface. Binding is also mediated by non- fimbrial adhesins (Jones et al. 1992, Antao et al. 2009).

Researchers are focusing in understanding bacterial pathogen adherence to host cells, by studying adhesin biogenesis, factors regulating adhesin expression and even efforts are being made to identify host receptors that are the targets of microbial adhesin factors (Finlay and Falkow 1997). Some of the important fimbriae and afimbrial adhesins for bacterial attachment and adhesion particularly for the ExPEC group of pathogens are discussed here.

2.1.1. Adhesins expressed by ExPEC

A variety of adhesins are expressed by ExPEC that prevent bacterial removal by host and target specific host epithelial cells to sustain itself within the host. The most important adhesins of ExPEC strains are fimbriae that are long surface protein structures and extending ~1 μ m out from the bacterial surface. A wide spectrum of fimbriae are expressed by ExPEC having different receptor specificities (Soto and Hultgren 1999, Klemm and Schembri 2000) providing bacteria the capacity to bind to different target molecules. Fimbrial adhesins recognize specific molecular motifs enabling the bacterium to target specific surfaces, such as a specific tissue in gut or bladder of human and animal hosts or respiratory tract. This phenomenon of tissue specificity is referred as tissue tropism that relies on specific interaction with receptor targets and specific tissue surfaces (Klemm et al. 2010). For example, type 1 fimbriae target uroplakins found in the bladder (Connell et al. 1996, Wu et al. 1996). Fimbrial adhesins expression functions in differential and coordinated expression manner, facilitating the bacteria to shift receptor target affinity and therefore change tissue preference (Holden and Gally 2004). For example: shifting expression of type 1 fimbriae to P fimbriae (Xia et al. 2000, Schembri and Klemm 2001b, Schembri et al. 2002, Snyder et al. 2005). This phenomenon of differential expression is important in ascending UTI because type 1 fimbriae target uroplakin receptors found in the bladder, whereas P fimbriae recognize the α -D-galactopyranosyl-(1-4)- β -D-galactopyranoside receptor found in the kidneys (Klemm et al. 2010) (Fig. 3).

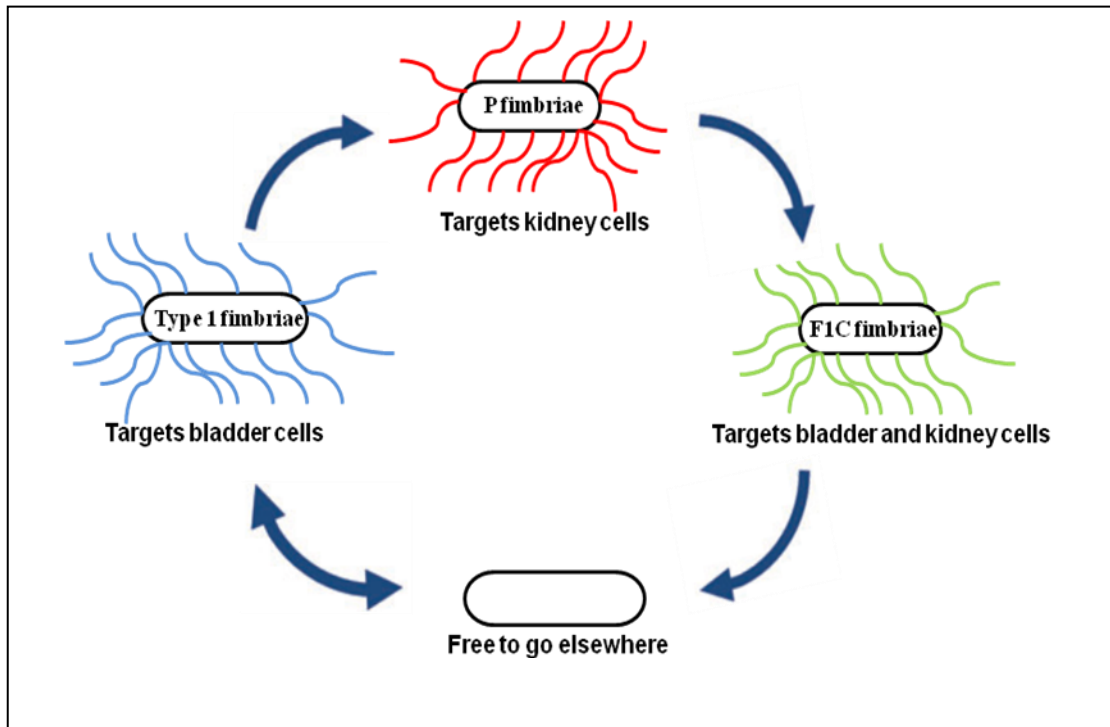


Fig.3 Model of tissue tropism as a function of differential fimbriae expression (modified from (Klemm et al. 2010)).

Different forms of adhesins present among ExPEC strains are Type 1 fimbriae (*fim*), P fimbriae or the pilus associated with pyelonephritis (*pap*), curli fibres (*csg*), S fimbriae or the sialic acid-specific fimbriae (*sfa*), F1C fimbriae (*foc*), Dr fimbriae (*dra*), afimbrial adhesins (*afa*), temperature-sensitive haemagglutinin (*tsh*) and other novel adhesin gene clusters that need to be properly characterized (Antao et al. 2009).

2.1.1.1 Type 1 fimbriae

Type 1 fimbriae (called also type 1 pili) are known to play an important role to promote bacterial adhesion, invasion and growth as a biofilm (Martinez et al. 2000, Schembri and Klemm 2001a). A bacterial cell possessing type 1 fimbriae has 100-500 fimbriae in number arranged peritrichously on the surface; each with a diameter of 7 nm and a length varying between 0.2 and 2 mm. The type 1 fimbrial organelle has a tubular structure with a ~2 nm diameter hollow core (Hahn et al. 2002). Type 1 fimbrial (*fim*) gene cluster contains four genes designated as *fimA*, *B*, *C* and *D* that are involved in the synthesis of the fimbriae (Klemm et al. 1985).

The bacteria shift periodically, the expression of type 1 fimbriae is phase variable fashion that decides bacterial state between a fimbriate and non-fimbriate form. Two regulatory *fim* genes, *fimB* and *fimE* control the phase variation of type 1 fimbriae in *E. coli* (Klemm 1986). Three additional genes *fimF*, *fimG* and *fimH* are involved in the adhesive property and longitudinal regulation of these structures (Klemm and Christiansen 1987). Binding to a variety of eukaryotic cells is conferred by Type 1 fimbriae by virtue of the capacity of the FimH adhesin to recognize mannosides (Krogfelt et al. 1990, Martinez et al. 2000).

The FimH adhesins from different strains show variation in their binding affinity towards defined oligomannose motifs and these differences in the binding capacity could be explained due to the alterations in the primary structure of FimH variants (Sokurenko et al. 1992, Sokurenko et al. 1994, Sokurenko et al. 1998). Receptor recognition profile can be affected by minor amino acid sequence alterations in FimH (Pouttu et al. 1999, Schembri et al. 2000). This variation in FimH adhesin enhances its binding capacity to targets such as laminin, collagen and fibronectin as well as to different mannose derivatives (Bouckaert et al. 2005). Allelic variations of FimH adhesin subunit are due to pathoadaptive mutations (Weissman et al. 2006, Weissman et al. 2007, Weissman et al. 2012) play an important role in tissue tropism determining fine sugar specificity of these fimbriae (Sokurenko et al. 1999, Oelschlaeger et al. 2002). Point mutation in FimH has been shown to enhance bacterial ability to adhere to CEACAM-expressing T84 intestinal epithelial cells (Dreux et al. 2013).

Type 1 fimbriae through its *fimH* adhesin play a significant role in the attachment of *E. coli* during infections, particularly UPEC infections (Eden and Hansson 1978). Studies on FimH revealed that FimH mediates biofilm formation (Pratt and Kolter 1998). Another feature associated with type 1 fimbriae is their ability to confer bacterial invasion into host cells, expression of type 1 fimbriae by UPEC promotes invasion of human bladder epithelial cells, while this phenotype was not observed in a *fimH*-negative mutant (Martinez et al. 2000).

2.1.1.2 P fimbriae

P fimbriae (encoded by the *pap* – pyelonephritis associated pili – operon) promote colonization by binding to the α -D-galactopyranosyl-(1-4)- β -D-galactopyranoside receptor epitopes of glycolipids on human erythrocytes of the P blood group and on uroepithelial cells and hence named P fimbriae and are mostly associated with disease related to the upper urinary tract

(Svanborg-Eden et al. 1987). The genes encoding the P pilus type are termed *pap* genes or pyelonephritis-associated pili genes since these are typical of strains isolated from human urinary tract infections (Hull et al. 1981).

The P fimbrial organelle comprises of multiple copies of major subunit protein (PapA) that is connected to a tip fibrillum comprising major (PapE) and minor (PapF, PapK, PapG) components. In animal models and in human infections the binding of P fimbriated UPEC triggers a proinflammatory cytokine interleukin (IL-6 and IL-8) production promoting the development of local inflammation response (Hedlund et al. 1999, Wullt et al. 2000, Godaly et al. 2000, Roberts et al. 2004).

The PapG adhesin which is located at the tip of P fimbriae helps to recognize their receptor targets (Kuehn et al. 1992). Several variants of the PapG adhesin are identified that recognize different isoreceptors and thus contribute to differences in tissue tropism by P fimbriated ExPEC strains (Lane and Mobley, 2007). The class II and class III PapG variants are mostly associated with human ExPEC strains (Klemm et al. 2010).

Nearly 80% of all pyelonephritis causing strains of *E. coli* express P fimbriae and recognition of galabiose receptor by PapG is thought to be a prerequisite for pyelonephritis (Thanassi et al. 1998). PapG is considered as a key determinant in promoting the virulence of *E. coli* in urinary tract infection (UTI) (Tewari et al. 1994). The P fimbriae are not only associated with Uropathogenic *E. coli* (UPEC) causing UTI, but are also related to newborn meningitis *E. coli* (NMEC) and avian pathogenic *E. coli* (APEC) (Guyer et al. 1998, Ewers et al. 2007, Simms and Mobley 2008).

2.1.1.3 F1C fimbriae

ExPEC strains of UTI origin also express F1C fimbriae (Usein et al. 2001, Snyder et al. 2005). F1C fimbriae are thin, 7-nm-wide and approximately 1 μ m long surface polymers. F1C fimbriae in their genetic organization and structural composition resemble type 1 fimbriae comprising of a major subunit protein (FocA), minor subunits (FocF and FocG) and a tip-located adhesin (FocH) (Klemm et al. 1994, Klemm et al. 1995).

F1C fimbriae have been implicated in the process of UTI and show binding to the epithelial cells of kidneys, ureters and bladder by attaching to galatosylceramide that acts as a specific target to FocH adhesin of F1C fimbriae (Khan et al. 2000, Backhed et al. 2002). F1C fimbriae contribute to the adhesive properties of UPEC strains and F1C-mediated bacterial attachment triggers innate immune system promoting human renal epithelial cells to produce pro-inflammatory cytokine, interleukin-8 in response to F1C-mediated attachment (Backhed et al. 2002). Like type 1 fimbriae, F1C fimbriae have recently been reported to enhance biofilm formation (Lasaro et al. 2009).

2.1.1.4 S fimbriae

S fimbriae are associated with those ExPEC strains causing neonatal meningitis and UTI. Most of the ExPEC isolates causing cystitis are found to express S fimbriae (Foxman et al. 1995, Bogyiova et al. 2002, Mabbett et al. 2009). S fimbriae were defined because of their receptor specificity and specific binding to sialyl galactosides. Morphologically, S fimbriae are similar to type 1 or P fimbriae of *E. coli* (Korhonen et al. 1984). Epithelial cells of the proximal and distal tubules, collecting ducts, and glomerulus, renal interstitium, and renal vascular endothelium are known to be the binding site for S fimbriae (Korhonen et al. 1986). S fimbriae also bind to the extracellular matrix components of fibronectin, laminin and sialoglycoproteins on brain microvascular endothelial cells which could be a possible explanation for migration of bacterial pathogens across physiological barriers (Wright and Hultgren 2006).

Earlier SfaA was considered to be the S fimbrial adhesin responsible for receptor recognition (Moch et al. 1987, Schmoll et al. 1989) and adherence properties (Prasadarao et al. 1993). However, based on recent studies of chaperone-usher mediated fimbrial biogenesis, SfaH represents the true adhesin of S fimbriae (Klemm et al. 2010). In a study on the prevalence of S fimbriae among ExPEC strains, it was observed that 50% UPEC, 24% NMEC and 9.2% APEC strains harbored the *sfa* genes (Ewers et al. 2007, Antao et al. 2009).

2.1.1.5 Curli fimbriae

Curli are thin, coiled, aggregative, amyloid-like fibres on the surface of *E. coli* and are considered to be the third category of *E. coli* surface organelles along with flagella and fimbriae, and are composed of a single type of subunit, curlin (Olsen et al. 1989). Curli

organelles appear as a tangled and amorphous matrix and are approximately 6-12 nm-wide fibres of various lengths extending 0.5-1 mm from the cell surface (Chapman et al. 2002). ExPEC strains associated with sepsis as well as some diarrheal *E. coli* pathotypes and avian pathogenic *E. coli* are known to express curli (Provence and Curtiss 1992, Olsen et al. 1993, Ben Nasr et al. 1996). Curli fibres are encoded on the *csg* (curlin subunit gene) gene cluster, consisting of two different operons, one which encodes the *csgB*, *csgA* and *csgC* genes, and second that encodes for *csgD*, *csgE* and *csgG* (Gophna et al. 2001). The production of the curli fibres requires expression of both operons (Olsen et al. 1998). *CsgD* is a transcriptional activator that is essentially required for the expression of the two curli fibre operons. Curli fibre expression is controlled by several elements, such as H-NS, RpoS and OmpR which results in a considerable reduction in curli fibres expression at temperatures higher than 30°C and at high osmolarity in most strains (Gophna et al. 2001). Curli expression is strongly associated with biofilm formation, adhesion to human proteins (e.g. laminin, fibronectin, plasminogen, major histocompatibility complex class I molecules) and invasion of eukaryotic cells (Olsen et al. 1989, Olsen et al. 1993, Ben Nasr et al. 1996, Olsen et al. 1998). Curli fibres are also found to play an important role for the internalization of bacteria causing avian septicemia as seen *in vitro* (Gophna et al. 2001).

2.1.2 Afimbrial adhesin expressed by ExPEC

Dr/Afa fimbriae

E. coli strains of diarrhoeagenic and UTI origin express Dr/Afa adhesins that comprise a family of fimbrial and afimbrial structures (Servin 2005). The structural assembly of genes coding for Dr/Afa adhesins consist of operons of five genes. These gene clusters are responsible for the biosynthesis of Afa adhesins belonging to the Afa/Dr family of adhesins and for the biosynthesis of invasins. AfaE-I adhesin involves five genes, *afaA*, *afaE*, *afaD*, *afaB*, and *afaC* (Labigne-Roussel et al. 1985). The AfaB, AfaC, and AfaE gene products are responsible for mannose-resistant haemagglutination (MRHA) and the *afaE* gene is identified as the structural gene encoding AfaE-I adhesin (Pham et al. 1997). The genetic organization of Dr adhesin operon also consists of five genes *draA*, *draB*, *draC*, *draD*, and *draE*. Four genes, *draA*, *draC*, *draD*, and *draE*, promote the expression of full, mannose-resistant haemagglutination (Nowicki et al. 1989). Dr/Afa adhesins recognize decay-accelerating factor (DAF) as a receptor which is a complement regulatory protein present on the surface of many human epithelial cells

(including epithelial cells of the urinary tract) (Labigne-Roussel et al. 1984, Medof et al. 1987). Some Dr/Afa adhesin variants are also known to mediate binding to type IV collagen (Berger et al. 2004).

2.2 Polysaccharide coatings

An important determinant in the virulence of extraintestinal pathogenic *E. coli* is the bacterial surface coatings, i.e. capsular polysaccharides (K-antigens), LPS, and O-polysaccharide moieties of LPS (O-antigens) (Russo et al. 1996). *E. coli* produces two types of capsular polysaccharide, designated as groups 1 and 2 that comprise a large number of distinct serotypes and are differentiated on the basis of chemical, physical, epidemiological, and microbiological characteristics. Most extraintestinal pathogenic isolates produce group 2 capsules (Jann and Jann 1987). Certain group 2 capsular serotypes contribute to pathogenesis in systemic models of infection (Cross 1990, Russo et al. 1994) and are more frequently present in strains that cause UTI. Group 2 capsules also contribute to urovirulence by their ability to prevent phagocytosis (Russo et al. 1996) and provide resistance to the bactericidal activity of serum (Leying et al. 1990, Russo et al. 1993). Studies show that when capsule-positive and capsule-negative strains were compared, the capsule-negative derivatives caused less bladder and renal colonization than their capsule-positive counterparts, thus further confirming that the capsular coatings contribute to the virulence properties of ExPEC (Russo et al. 1996). O-specific antigens are also considered as possible virulence determinants for systemic infection and in the pathogenesis of UTI (Russo et al. 1993). O-specific antigens are more prevalent in the strains that cause pyelonephritis and symptomatic cystitis and also confer serum resistance among these strains (Russo et al. 1996). There are 180 different O serogroups and more than 80 K-antigens in *E. coli* (Reisner et al. 2003, Stenutz et al. 2006) however, 75% of the urinary tract infections are caused by *E. coli* from a small number of O serogroups – O4, O6, O14, O22, O75 and O83 (Stenutz et al. 2006).

2.3 Lipopolysaccharide (LPS)

Lipopolysaccharide (LPS), also known as endotoxin, is an important component of the outer membrane of Gram negative bacteria and is main bacterial factor for development of endotoxemia as a result of systemic immune response leading to the lethal shock condition of the host when it encounters the bacterial pathogen (Zimecki et al. 2004). LPS consists of three

parts: lipid A, forming the toxic component; the core region, that could be divided into an inner and an outer part; and finally the O-antigen polysaccharide, which is specific for each serogroup (Brade 1999). In addition, many extraintestinal *E. coli* strains can synthesize colonic acid, which is a main component of the polysaccharide coatings (Reisner et al. 2003) and studies show that bacterial coatings could provide increased serum resistance and facilitate ExPEC to avoid clearance by phagocytosis (Suerbaum et al. 1994, Rizvi and Kumar 2003).

In UTI cases, the O-polysaccharide moiety of the UPEC has an important regulatory function (Svanborg et al. 2001, Fischer et al. 2006) and promotes signaling through Toll-like receptor -4 (TLR4) of the urothelial and inflammatory cells leading to the production of pro-inflammatory mediators (e.g. cytokines and chemokines (Austin et al. 2003, Backhed et al. 2003). As a result of inflammation the mucosal barrier breaks allowing UPEC to gain access into the underlying tissues (Fischer et al. 2006). In APEC, O78- antigen has been shown to be responsible for increased bacterial resistance against serum (Mellata et al. 2003).

2.4 Toxins and proteases

Bacterial toxins (endotoxins and exotoxins) and proteases are considered important attributes that define virulence properties among ExPEC. They play a promising role in different biological activities carried by bacterial pathogens that include cell adhesion, iron accumulation, and cell invasion through cell lysis and disruption of the mucin layer and the epithelium, as well as modulation and induction of the cell cycle, inflammatory reactions, and apoptosis (Nougayrede et al. 2005).

2.4.1 Haemolysin

A major virulence factor for some of the pathogenic strains related to extraintestinal disease like urinary tract infections, peritonitis, meningitis and septicemia is the production of haemolysin (Cavalieri et al. 1984). Haemolysin is called so as it lyses red blood cells and acts mainly by attacking the immune system cells of the host, severely impairing their function (Coote 1995). Haemolysin activity is a multi-step process that involves release of iron from erythrocytes, disruption of phagocytic function, and ultimately causing direct toxicity and damage to host tissues (Johnson 1991).

E. coli strains produce different types of haemolysins, mainly secreted (alpha haemolysin), bound (beta haemolysin) and gamma haemolysin that are produced by nalidixic resistant strains. Among all the haemolysins, alpha haemolysin is the major virulence factor for causing several extraintestinal infections (Cavalieri et al. 1984). It is a representative prototype of a family of homologous cytolytic protein toxins, namely RTX (repeat in toxin) family that are secreted by Gram-negative bacteria, e.g. *E. coli*. (Murase et al. 2012).

In an *E. coli* isolate that is related to urinary tract infections a coding region of 7 kb was identified and was found to be responsible for haemolysin synthesis (Noegel et al. 1981). Four genes, *hlyA*, *hlyB*, *hlyC*, and *hlyD* are involved in extracellular haemolysin activity (Goebel and Hedgpeth 1982). *HlyA* gene encodes the structural haemolysin protein activated by intracellular protein HlyC, thereby facilitating binding of haemolysin with erythrocytes causing hemolytic activity (Felmlee et al. 1985, Oropeza-Wekerle et al. 1989). HlyB is involved in providing energy during haemolysin secretion as it is an energy dependent process, whereas HlyD is required for the release of haemolysin from the outer membrane (Wagner et al. 1988, Oropeza-Wekerle et al. 1989).

Alpha haemolysin causes osmotic lysis of erythrocytes by pore forming activity and is cytotoxic to all mammals (Rennie and Arbuthnott 1974). Lysis of erythrocytes is a calcium dependent process as the calcium modifies haemolysin aggregates, making them haemolytically competent (Short and Kurtz 1971). The UPEC HylA has a dual physiological function; at higher concentrations, this toxin has pore forming activity on plasma membrane showing its cytolytic behavior (Soloaga et al. 1999), and at lower concentrations interferes with the Ca²⁺ dependent signaling pathways and hence modulating inflammatory responses of the target cells (Uhlen et al. 2000, Soderblom et al. 2002, Oxhamre et al. 2005).

2.4.2 Cytotoxic necrotizing factor (CNF Toxin)

Cytotoxic necrotizing factors (CNFs) are *E. coli* associated protein toxins that cause necrosis and induce formation of multi-nucleation of cells and enlargement of eukaryotic cells in cell culture (Caprioli et al. 1987). Forty percent of UPEC strains produce CNF-1 and this is also produced by 5%-30% of *E. coli* strains that cause diarrhea. CNF-1 acts via activation of Rho, which is small guanosine triphosphate (GTP) binding protein that helps CNF-1 affecting epithelial cells, endothelial cells, or monocyte-macrophages (Hofman et al. 2000). *E. coli*

isolates that cause UTIs in humans and produce haemolysin also synthesize CNF-1 and this phenotypic association is due to the fact that the determinants responsible for the production of these two toxins are in a state of genetic linkage, on the chromosome of uropathogenic *E. coli* (Falbo et al. 1992, Blum et al. 1995).

Cytotoxic necrotizing factors 1 and 2 (encoded by *cnf* genes; CNFs) are produced by intestinal and extraintestinal pathogenic *E. coli* and are lethal to a wide variety of eukaryotic cells (Hoffmann and Schmidt 2004, Landraud et al. 2004). The CNF toxins consist of two domains: the cell-binding domain and the catalytic domain (Lemichez et al. 1997). The host cell Rho GTPases are activated by CNF thus influencing the regulatory pathways of the actin cytoskeleton, transcription, cell transformation, cell proliferation, and apoptosis (Oswald et al. 1994, Mills et al. 2000). All these events increase the phagocytic activity of the epithelial cells enabling the CNF-producing *E. coli* to cross the epithelial barrier (Landraud et al. 2004). In addition, CNFs influence the secretion of inflammatory mediators by modulating and impairing immune responses (Malorni et al. 2003, Munro et al. 2004). CNF-1 is localized in the cytoplasm of UPEC bacterial cells. CNF-1 is also associated to outer membrane vesicles that help the bacteria to secrete and deliver CNF-1 to the environment and to cause infection to the host tissue.

2.4.3 SPATE (serine protease autotransporters of Enterobacteriaceae) family

Virulent strains of *E. coli* display specific phenotypic traits and secrete specific proteins that may contribute to their pathogenesis. One of these protein is the secreted autotransporter toxin (Sat) which is a serine protease and is found predominantly in uropathogenic strains of *E. coli*. Sat belongs to a subgroup of autotransporters classified as the SPATE (serine protease autotransporters of Enterobacteriaceae) family. Some SPATE proteins are: Sat that belongs to uropathogenic *E. coli*; Tsh that is prevalent in avian *E. coli* (Guyer et al. 2000).

SPATE autotransporters are identified by the presence of a serine protease active site motif that is involved in phenotypic functions (e.g., adhesin, invasins, protease, or cytotoxin). Sat is known to exhibit cytopathic activity on HEp-2 cells (Guyer et al. 2000) and is also a vacuolating cytotoxin, causing vacuolation within the cytoplasm of mammalian bladder and kidney cells (Guyer et al. 2002).

Another serine protease autotransporter is Tsh, that is identified as a temperature-sensitive protein causing haemagglutination of chicken red blood cells and cleavage of hemoglobin (Provence and Curtiss 1994). Tsh is associated with avian pathogenic *E. coli* isolates that are responsible for the development of lesions and fibrin deposition in the air sacs of chickens (Dozois et al. 2000). A hemoglobin protease identical to Tsh protein, with only two amino acid differences (Q209-K209 and A842-T842), named Hbp was characterized from a human pathogenic *E. coli* that specifically degrades human hemoglobin and binds to the released heme, suggesting its role in heme acquisition (Otto et al. 1998).

The UPEC PicU (*picU*) is another serine protease, homologous to Pic (protein involved in intestinal colonization) of *Shigella* spp. and enteroaggregative *E. coli*. Pic proteases are involved in mucinase activity and cleave human complement (Henderson et al. 1999, Parham et al. 2004). OmpT (*ompT*) outer membrane endopeptidase of ExPEC is also a serine protease, similar in structure to plasminogen activator of *Y. pestis* and is capable of degrading antimicrobial peptides that are secreted by the epithelial cells and macrophages (Grodberg et al. 1988, Kukkonen and Korhonen 2004).

Thus, all phenotypes associated with the SPATE proteins are involved in the mechanisms that enable bacteria to damage the host or avoid an immune response.

2.5 Iron acquisition systems

Iron is an essential cofactor for bacterial metabolism, survival and multiplication in certain ecological niches and depends on the ability of these organisms to scavenge these essential nutrients. Acquisition of iron is an important prerequisite for pathogenic bacteria to carry out infectious process, but sufficient amount of iron is also required by non-pathogenic *E. coli* (Mietzner and Morse 1994). Iron plays a vital role in cellular processes like energy generation, DNA replication, oxygen and electron transport, metabolism of peroxidases and protection against oxidative stress (Skaar 2010). Within the mammalian host, free ionic iron concentration is extremely low (10^{-18} M) and the availability of free iron is further reduced by iron chelating host proteins like, albumin, heme, hemoglobin, ferritin, transferrin, and lactoferrin (Bullen et al. 2005). This depletion of free iron for microbes is a kind of defense mechanism adopted by the host against bacterial infections preventing bacterial growth and is termed nutritional immunity (Skaar 2010). Therefore, in order to cope up with the scarcity of iron, pathogenic *E. coli* have

come up with several effective iron uptake systems and these specialized ability of iron accumulation significantly contributes to the virulence of these strains (Bullen et al. 2005). The iron concentration in *E. coli* ranges from $\sim 10^5$ to 10^6 atoms per cell depending on growth conditions (Abdul-Tehrani et al. 1999, Andrews et al. 2003).

Pathogenic bacteria, including ExPEC and more specifically UPEC, swipe iron from their host by expressing an iron acquisition system that utilize siderophores to scavenge iron from the environment and making it available to the bacterial cell. Siderophores are low molecular weight compounds that have a high affinity for ferric (Fe^{3+}) iron, thus competing with host proteins that chelate iron. Iron-bound siderophores are retrieved by bacteria through receptors that facilitate the transport of siderophore-iron complexes through the bacterial membrane and thus, releasing iron into the cytosol (Wiles et al. 2008).

A common siderophore known as enterobactin is encoded by the *ent* gene cluster and its receptor FepA is produced by pathogenic strains like UPEC and non-pathogenic *E. coli* (Raymond et al. 2003). Enterobactin effectively chelates iron from the host by outcompeting host transferrin for iron binding and hence making iron available to bacteria further facilitates bacteria to colonize poor iron environments like urinary tract (Demir and Kaleli 2004).

In order to counteract enterobactin-mediated iron scavenging by pathogens, the host produces a host protein lipocalin 2 (also known as neutrophil gelatinase-associated lipocalin, siderocalin, 24p3, or uterocalin) that specifically binds to enterobactin and stops iron scavenging (Goetz et al. 2002). Lipocalin 2 is expressed and released by neutrophils that are the major immune effector cells employed to the sites of infection within the urinary tract (Reigstad et al. 2007). But, ExPEC and UPEC in particular, express different iron acquisition systems, including the siderophores salmochelin and yersiniabactin. Interestingly, salmochelins are variants of enterobactin that gets modified by glycosylation via the action of a glucosyltransferase encoded within the *iroA* gene cluster (Bister et al. 2004, Smith 2007) and this modification prevents salmochelins recognition and sequestration by lipocalin 2 that give *iroA*-positive bacteria a distinguish advantage within the host thereby fulfilling their iron requirement (Fischbach et al. 2006). The *iroA* gene cluster is carried within PAIs by many UPEC isolates. However, iron-bound salmochelin is not recognized by the normal enterobactin receptor FepA, and so the *iroA* gene cluster encodes another receptor IroN, which recognizes iron-bound salmochelin and

transports it into the bacterial cytosol. IroN mediates the uptake of salmochelin and other catecholate siderophores (e.g. enterobactin and dihydrobenzoic acid) and contributes to the virulence of UPEC and NMEC causing UTI and neonatal meningitis respectively (Hantke et al. 2003, Negre et al. 2004).

Another siderophore is aerobactin which is an important hydroxamate iron uptake system of pathogenic and non-pathogenic *E. coli* and significantly contributes to the virulence of ExPEC (Torres et al. 2001). The aerobactin gene cluster (*iuc*) and ferri-aerobactin receptor (*iutA*) are often located on ColV plasmids and were first identified in enteroinvasive *E. coli*. This Iuc-IutA iron uptake system is well distributed throughout the intestinal and extraintestinal pathogenic *E. coli* strains, and virulence of these strains is determined by the expression of these iron uptake system (Demir and Kaleli 2004).

Yersiniabactin is another known catecholate siderophore that is expressed by ExPEC strains and was first described in *Yersinia pestis* (Fetherston and Perry 1994). Among the ExPEC strains, a conserved chromosomal gene island denoted as the High Pathogenicity Island encodes for the yersiniabactin biosynthetic gene cluster (*irp*) and the yersiniabactin receptor FyuA (Carniel et al. 1996). The *irp-fyuA* gene cluster expression by ExPEC contributes to the virulence of these strains (Koczura and Kaznowski 2003).

2.6 Serum resistance

Serum has a lethal effect on Gram negative bacteria and plays an important role in the host defense mechanism against foreign invasion. The bactericidal activity of the serum is due to the complement system and this lethal property of serum is lost when the complement is inactivated or depleted. Bacterial resistance to the bactericidal effect of serum is shown by many pathogenic strains and particularly by invasive strains and those causing UTIs (Taylor 1983, Burke et al. 1990). Extraintestinal human pathogenic and avian *E. coli* strains possess many virulence factors, such as K1 capsules, lipopolysaccharide (LPS), outer membrane proteins (OMPs), including OmpA and the ColV plasmid-encoded proteins TraT and Iss, that contribute to the bacterial resistance against the complement system of the host (Gyles 1994).

Capsular polysaccharides are a coating on bacterial cells and are important in the virulence of many bacterial pathogens by providing resistance to phagocytosis and protection against

complement mediated killing (Buckles et al. 2009). Extraintestinal pathogenic *E. coli* strains have thin, acidic, thermostable and highly anionic capsule (Jann and Jann 1983). Strains with K1 capsule are less prone to phagocytosis than the noncapsulated strains but heat-inactivated capsular polysaccharide strains undergo phagocytosis (Jann and Jann 1983, Allen et al. 1987). The negative charge and the hydrophilicity of the K1 polysaccharide are responsible for the antiphagocytic property of K1 capsule (Eisenstein and Jones 1988). Capsular polysaccharide also blocks opsonization by preventing complement activity towards bacterial cells (Johnson 1991). The capsular polysaccharide enhances bacterial survival in serum making it an important determinant for serum resistance (Taylor 1983). Human strains, involved in urinary tract infections and those causing cystitis and pyelonephritis include certain K types, like K1, K2, K3, K5, K12, K13, K20, and K51 (Johnson 1991). Most of the strains that cause meningitis also possess K1 capsule (Czirok et al. 1986).

Bacterial resistance to the killing effect of serum also results from O-polysaccharide side chains and other surface proteins. It has been well documented that the smooth strains are more resistant to serum than the rough strains and that this resistance is directly proportional to the amount of lipopolysaccharide present on the cell. The O polysaccharide is a component of smooth-type lipopolysaccharide. Serum-resistant strains with abundant O-polysaccharide are shielded against the action of complement as compared to the serum sensitive strains. O-antigen polysaccharide side chains protect against complement lysis by blocking the complement to act on the sensitive target sites on the membrane of bacterial cell thus, preventing it from the killing action of complement (Taylor 1983, Goldman et al. 1984).

Outer membrane protein A (OmpA) is one of the most abundant proteins of bacterial cell envelope (Klose et al. 1993). The outer membrane protein A (OmpA) contribute to the bacterial resistance towards the lethal effect of serum complement system. Studies have shown that the *ompA* mutant is more sensitive to the bactericidal effect of human serum by complement activation as compared to the strains having OmpA. This protein serves to stabilize the outer membrane of bacterial cell, making it more resistant to complement action, thus contributing to the pathogenesis of *E. coli* infections (Weiser and Gotschlich 1991).

There are certain plasmids that might confer increased serum resistance of the host strain (Taylor and Hughes 1978). ColV plasmids enhance *E. coli* virulence properties like

complement and phagocytosis resistance enabling *E. coli* strains better survival advantage within the host. ColV plasmids suggest pathogenicity of an organism and are frequently found in isolates from extraintestinal infections in animals and humans (Fernandez-Beros et al. 1990). In human *E. coli* isolates ColV plasmids mediate serum resistance by suppressing killing action of complement by producing Iss (increased serum survival) protein (Johnson 1991).

Iss occurs as 10-11 kDa protein in the bacterial outer membrane and there are approximately 2000 molecules of Iss per bacterial cell. Iss protein has anti-complementary effect and is involved in surface exclusion. Avian pathogenic *E. coli* associated with systemic *E. coli* infection were found to carry *iss* more frequently as compared to the healthy birds and *iss* is considered to be more relevant in complement resistance of avian *E. coli* compared with human *E. coli* (Nolan et al. 2003)

Another outer membrane protein is TraT that is associated with the ColV plasmid and confers serum resistance to bacterial cell by interfering complement mediated killing mechanism of host serum. TraT interferes complement action by surface exclusion and is also produced by *E. coli* strains that carry F-like plasmids such as R6-5 and R100 which promote production of the TraT protein (Nilius and Savage 1984, Johnson 1991). The cloned *traT* gene was observed to increase serum resistance in uncapsulated strains (Aguero et al. 1984).

3. Diseases that are related to ExPEC

3.1 Urinary tract infections

Urinary tract infections are caused by uropathogenic *E. coli* (UPEC) and this pathotype accounts for 80% of all UTIs causing cystitis in urinary bladder, pyelonephritis in kidneys, bacteriuria when present in urine and prostatitis when present in prostate. Individuals that are at higher risk of acquiring UTIs are pregnant women, elderly patients, infants, and patients with spinal cord injuries, diabetes, multiple sclerosis, urinary catheters, and HIV/AIDS (Foxman 2002). Most of the strains that are responsible for UTIs belong to phylogenetic group B₂ (Johnson et al. 2005b) and to a lesser extend to phylogenetic group D (Bonacorsi et al. 2005, Bidet et al. 2005).

There are several virulence associated genes among *E. coli* isolates causing UTIs that serve as an important attribute during the course of infection and include *iuc* (aerobactin) *kpsMTII* (KII

capsule), *cnf-1*, *dra* (Dr-binding adhesins), *hly*, *ompT*, *papGI* (P pili fimbriae class I), *papGII* (P pili fimbriae class II), *papGIII*, *sfa*, and *fim* (Marrs et al. 2002, Marrs et al. 2005).

Different stages of pathogenesis leading to the development of UTIs include bacterial adherence to the host tissue, colonization within the host, avoidance of host defense mechanisms and causing damage to host tissues (Connell et al. 1996). The entry of uropathogens into the urinary tract requires adhesion to the uroepithelium and is mediated by fimbrial adhesin FimH of Type 1 fimbriae that have been studied a lot and are known to play a critical role in the early stage of bacterial infection mediating bacterial adherence to the host tissue (Gunther et al. 2002). However, P fimbriae are also known to be associated with *E. coli* causing UTI (Johnson et al. 2005a).

Once the pathogen attains attachment to the host tissue they start colonizing within the host and in order to survive and grow within the host, bacteria need to scavenge iron from their surroundings. Bacteria employ different siderophores in order to fulfill their iron requirement that trap and deliver iron from the surrounding. The *iroN* gene encodes for the receptor of enterobactin siderophore that captures iron and makes it available to the uropathogens thus helping in their colonization. Likewise other siderophores, such as aerobactin also quench iron and help uropathogens to colonize the urinary tract (Torres et al. 2001, Russo et al. 2002).

In order to successfully colonize host tissues, pathogens move and localize themselves to a better niche. Uropathogens possess a *fliC* gene that encodes flagellin, the structural subunit of flagella which allows for this movement and thus contributes in the infection mode of UPEC in the urinary tract (Wright et al. 2005).

After colonization, escaping host defense mechanisms is on priority and uropathogens employ a gene *surA* that has been studied to play an important role in bacteria's invasion into and propagation in the host cells thereby preventing uropathogens from host defense counteraction (Justice et al. 2006). Another gene, *oxyR* is also important as it protects uropathogens from the oxidative stress defense mechanism that host recruits against bacterial pathogens (Johnson et al. 2006).

Once the uropathogens establish themselves within the host they secrete toxins to damage host, in order to facilitate better survival environment. Three toxins are produced by uropathogens

and the genes that encode these toxins, which account for important virulence factors of uropathogens, are *cnf-1*, *hly*, and *sat* (Rippere-Lampe et al. 2001, Guyer et al. 2002, Nagy G 2006). The cytotoxic necrotizing factor is encoded by *cnf-1* gene and protects the bacteria against the killing action of neutrophils. The haemolysin is encoded by *hly* gene and forms pores in host cells leading to their destruction causing apoptosis, while the *sat* gene encodes a cytotoxin that forms vacuoles in host cells, most prominently in human kidney cell lines (Smith et al. 2007).

3.2 Neonatal meningitis

Neonatal meningitis associated *E. coli* (NMEC) is the most common cause of Gram-negative associated neonatal meningitis that can be fatal and leads to severe neurological defects in many survivors (Unhanand et al. 1993). The infection can be acquired from the mother during birth, or through *E. coli* infections of umbilici, the infant's upper respiratory or intestinal tract. NMEC-induced neonatal meningitis involves several steps like: bacteremia i.e. the binding of bacteria to the surface of neonatal brain microvascular endothelial cells (BMECs), followed by bacterial invasion of BMECs and later on invasion of the meninges (membranes that surround the brain and spinal cord) and the central nervous system (CNS) (Smith et al. 2007). Nearly 80% of the NMEC strains that cause meningitis carry a K1 capsular antigen. The most frequent serotype among NMEC strains is O18:K1 and the majority of the strains belong to the phylogenetic group B2 while other serotype like O1:K1 and strains belonging to phylogenetic group D are also found among NMEC strains (Johnson et al. 2002b, Mora et al. 2009, Cortes et al. 2010).

The presence of Outer membrane protein A (OmpA), K1 capsular polysaccharide antigen, and O-lipopolysaccharide (O-LPS) are some of the important attributes that protect NMECs against serum complement system, opsonization and phagocytic killing by host defense mechanisms thereby providing better chances of survival and multiplication of NMECs in the circulatory system resulting in the condition of bacteremia (Xie et al. 2004). ExPEC neonatal meningitis strains may harbor many virulence associated genes on pathogenicity associated islands (PAIs) for example, the genes *sfaS*, *ibeA*, and *cnf-1* are located on PAIs and are involved in facilitating blood-brain barrier penetration by the microorganisms (Bonacorsi et al. 2003). Bacterial invasins like IbeA, IbeB and IbeC proteins, AslA contribute BMEC invasion by K1-producing

E. coli. Invasion of BMECs involves rearrangement of the actin cytoskeleton by NMEC and here, virulence factors like OmpA and CNF-1 are involved (Kim 2001).

K1-bearing NMEC invade BMECs and internalize themselves within BMEC membrane-bound vacuoles. The K1 capsule inhibits vacuole maturation as a result of which lysosomes do not fuse with the vacuoles containing bacteria, facilitating intracellular bacterial survival and transversal to human BMECs as live bacteria. Therefore, *E. coli* invade BMECs via transcellular process. Once the NMEC traverse the BMECs, they invade the meninges and CNS and induce the release of proinflammatory compounds (cytokines, chemokines, reactive oxygen species, nitric oxide) leading to an increase in the permeability of blood-brain barrier and pleocytosis (increase in leukocytes in the spinal fluid) resulting in brain edema and increase in intracranial pressure finally lead to meningitis and neuronal injury (Kim 2003).

3.3 Sepsis

The clinical condition when bacteria are present in the circulatory system is termed as bacteremia. At the later stage, when the bacteria start multiplying themselves, bacteremia progresses towards septicemia. Sepsis is a life-threatening severe illness condition due to the bacterial infection of the blood stream and is alternatively known as systemic inflammatory response syndrome (SIRS). In severe sepsis condition there is dysfunction and failure of organ (Andreoli 1997, Annane et al. 2005). The most common causative pathogen for sepsis is *E. coli* and the pathotype is termed as sepsis associated pathogenic *E. coli* (SePEC), while other *E. coli* pathovars might likewise be able to cause septicemia under certain conditions. Most of the sepsis causing *E. coli* belongs to phylogenetic group B₂ (Johnson et al. 2002a).

Septicemia in humans could also result from the later stages of UTI or respiratory diseases. Sepsis occurs as a result of microbial infection originating from the kidneys (UTI), bowel (peritonitis), skin (cellulitis), or lungs (pneumonia) and from other body locations. Newborns, immunocompromised persons and elderly patients are at higher risk of bacterial sepsis (Stoll et al 2005). Sepsis occurring in the newborn infants is known as neonatal sepsis and may be due bacterial infection acquired during the passage of birth canal or from hospital or home environment (Schrag and Schuchat 2005). Septicemia is an important disease related to the farm animals where it causes colisepticemia in livestock animals and especially avian colisepticemia leading to significant economic loss (Mokady et al. 2005).

Many *E. coli* strains that are associated with sepsis belong to O2 and O78 serogroups (Ananias and Yano 2008). Mokady et al. (2005) determined virulence factors among these *E. coli* strains which refer to adhesins (type 1 pili, curli, and P pili), iron uptake systems (aerobactin, yersiniabactin, and Iron receptor) and serum resistance (Mokady et al. 2005). An important virulence factor i.e. the type III secretion system has been identified among septicemic ExPEC as well (Ideses et al. 2005). Plasmid-encoded Colicin V (ColV plasmid) is carried by many of the septicemic strains but not all however, it has been suggested that the presence of ColV gene together with other plasmid-located virulence genes are responsible for full virulence (Tivendale et al. 2009). Another important virulence feature is the secretion of OmpA by septicemic *E. coli* that nullifies the bactericidal effect of neutrophils elastase by binding it (Mokady et al. 2005).

3.4 Avian systemic infection

ExPEC strains that are the causative agents of avian colibacillosis in poultry are termed as avian pathogenic *E. coli* (APEC). Colibacillosis disease is a major threat to poultry industries causing huge economic loss. The most common complications among poultry birds associated with colibacillosis are perihepatitis, airsacculitis and pericarditis (Dziva and Stevens 2008). Initially, APEC enters the respiratory tract and colonizes the air sacs. Subsequent infection is followed as aerosacculitis, later leading to septicemia, pneumonia, pericarditis, perihepatitis, and peritonitis and finally results in the death of the bird (Dho-Moulin and Fairbrother 1999, Li et al. 2005). Three serotypes O1, O2 and O78 are predominantly present in APEC, however other serotypes, such as O93 and O92 were also observed among APEC strains (Wang et al. 2010). Different virulence factors associated with APEC strains include K1 capsule, type 1 and P fimbriae, curli, aerobactin, and temperature-sensitive haemagglutinin (Tsh) (Janßen et al. 2001, Ewers et al. 2005, Johnson et al. 2008c). Another virulence mechanism for the pathogenesis of APEC strains is serum resistance and is mostly related with strains causing septicemia in chickens (Ewers et al. 2009, Li et al. 2011).

Many of the virulence attributes that contribute to the complement system resistance of APEC strains are also common in human and animal ExPEC strains causing neonatal meningitis and UTIs in humans and UTIs in animals (Schouler et al. 2004, Ron 2006). Studies have also shown that UPEC and APEC have similarities in their serotype and phylogenetic groups. The most

common serotypes among UPEC are also shared by APEC. For example the O2 serotype commonly found in UPEC, is also a common occurring serotype of APEC, suggesting that poultry may be the carrier for *E. coli* capable of causing urinary tract infections in humans and further hypothesizing that avian ExPEC strains could have zoonotic nature (Johnson et al. 2007, Ewers et al. 2007, Moulin-Schouleur M 2007, Mora et al. 2009). Rodriguez-Siek et al. have suggested that certain APEC strains are potentially competent to infect humans and can even act as a reservoir for the virulence genes for ExPEC (Rodriguez-Siek et al. 2005).

4. Zoonotic nature of ExPEC

As defined by WHO (World Health Organization) zoonosis is any disease or infection that is naturally transmissible from vertebrate animals to humans. Bacterial species that cause infections in animals and are potentially capable of causing infections, also in humans or vice versa are known to have zoonotic nature. The origin of ExPEC infecting humans is still unclear, but it is possible that the ExPEC from animals could cause infections in humans as animals are reservoir for ExPEC. ExPEC infections are regularly observed in poultry and in livestock animals. Also the companion animals (dogs and cats) suffer from infections caused by ExPEC (Ewers et al. 2007, Belanger et al. 2011).

The commonalities in phylogenetic origins of the strains isolated from humans and animals, as well as their genome flexibility are of serious concern making ExPEC potentially competent to cause zoonosis (Brzuszkiewicz et al. 2009) and providing ExPEC the capability to cross the species barrier and effectively colonize humans as well as companion animals such as dogs and cats (Johnson et al. 2009). A number of human-associated ExPEC virulence genes are found among the canine ExPEC strains, indicating that dogs may be the reservoir of human ExPEC strains. For example, both canine and human ExPEC strains express *papG* allele III and there was no difference in the peptide sequences among the *E. coli* isolates from humans and dogs indicating that the *papG* allele III sequences were highly homologous among these strains (Johnson et al. 2000). The common characteristics between human and canine ExPEC strains points out the fact that humans may acquire ExPEC strains from dogs. Since there is similarity in virulence genes of human and dog ExPEC strains there could exist similar pathogenic mechanisms. If humans are colonized with canine-derived ExPEC strains and antimicrobials are used in veterinary practice, it could lead to the selection of new antimicrobial resistant human

pathogens (Smith et al. 2007, Ewers et al. 2011). As the human and animal strains could not be differentiated according to their specific host, virulence gene and serotype profiles, or pulsed field gel electrophoresis profiles certain pathogenic lineages of ExPEC might cause disease in both animals and humans suggesting the possibility of cross-species infection (Johnson et al. 2001a).

Poultry is also considered to be a reservoir for ExPEC that may introduce these strains into the environment and later cause infection to other hosts. The ExPEC strains of human and avian origin show similarities that suggest that the avian strains potentially have zoonotic properties (Johnson et al. 2007, Ewers et al. 2007, Schouleur et al. 2007, Mora et al. 2009). Several studies have indicated a commonality of virulence factors in human and animal strains of ExPEC (Smith et al. 2007) and that APEC strains have the same serotypes like that of the strains that are responsible for extraintestinal infection in humans (Rodriguez-Siek et al. 2005, Moulin-Schouleur M 2007, Mora et al. 2009). It was also found that there exists a genomic similarity in the first fully sequenced APEC isolate and several human ExPEC strains (Johnson et al. 2007). Based on their screening of a large collection of APEC, UPEC, and NMEC strains for virulence features, O-types, and phylogenetic background Ewers et al. (2007) speculated that poultry may act not only as a carrier but also as a reservoir of virulence-associated genes for human ExPEC strains that may transform commensal *E. coli* of human gut into extraintestinal UPEC and NMEC strains (Rodriguez-Siek et al. 2005, Schouleur et al. 2006, Ewers et al. 2007, Johnson et al. 2008d).

5. Multilocus sequence typing (MLST)

In surveillance the ability to identify infectious strains accurately is most important. Molecular typing methods are used to address two important queries; (1) are the isolates found from a localized outbreak of disease the same or different and (2) are strains causing disease in one geographic area related to those isolated elsewhere world-wide (Maiden et al. 1998). Sequencing-based genotyping methods like single-locus sequence typing (SLST) and multilocus sequence typing (MLST) are more widely being used in epidemiological studies of infectious diseases, including UTIs, meningitis etc. In SLST sequence data for specific loci (e.g. involved in virulence, drug resistance, pathogenicity, etc.) from different strains of the same species for variability in a specific gene, such as single-nucleotide polymorphisms and areas

with repetitive sequence is being studied. Thus, making it a promising sequence-based tool for epidemiologic analysis (Singh et al. 2006).

Another sequence-based method is multilocus sequence typing (MLST), a technique first introduced by (Maiden et al. 1998) that provides reproducible, comparable and transferable results among different laboratories. MLST is based on sequencing ‘housekeeping’ genes, which are under stabilizing selection. MLST utilizes more representative and larger portion of the genome than SLST. MLST utilizes the comparison of the nucleotide sequences of internal 400- to 500-bp regions of a series of housekeeping genes (typically seven or more) present in all isolates of a particular species. Distinct alleles are classified based on genetic polymorphism in sequences for each gene fragment. Each isolate is defined by the alleles at each of the sequenced housekeeping loci that together form the allelic profile or sequence type (ST), as there are many potential alleles at each of the loci, it is unlikely that identical allelic profiles will occur by chance. Therefore, isolates with the same allelic profile are assigned as members of the same clone (Singh et al. 2006). Thus, MLST helps in better understanding the genetic interrelationships in bacterial population.

6. ExPEC related to sequence type 95 (ST95)

As discussed, ExPEC frequently cause urinary tract infections (UTIs), septicemia, meningitis, respiratory tract and soft tissue infection etc. in human and in different animal hosts. In recent years great efforts have been made to discern the group of ExPEC with respect to clinical implications and host origin, but still a precise definition and thus a sound risk assessment of ExPEC strains, in particular with regard to their zoonotic potential, is not found. The genetic background of *E. coli* is determined by Multilocus sequence typing (MLST) which assigns isolates to different sequence types (STs). Interestingly, most of the sequence types have mixed groups of strains from different sources, including humans, birds, dogs, cats, cattle, horses, swine, etc.

In contrast, one particular prominent ST of ExPEC is ST95 that accumulated ExPEC strains from human and avian origin of particularly high virulence possessing a great set of virulence associated genes (Ewers, Wieler et al., unpublished data). ST95 strains mostly originate from humans and birds indicating a putative host specificity of these strains. Hence, ST95 resembles a group of strains which was termed by our group as “Human-avian” complex, as according to

the publically available database (<http://mlst.ucc.ie/mlst/mlst/dbs/Ecoli/> as observed on April 2013) it currently contains strains mostly from human and avian origin with one canine strain thus, supporting some kind of host specific nature ST95 strains.

Currently it is unclear whether all ST95 strains are zoonotic, or whether some strains are specific to one particular host (Schouleur et al. 2007, Johnson et al. 2008d, Tivendale et al. 2010, Bauchart et al. 2010), namely the human or an avian host. We hypothesized that the host-specificity may be caused due to the genetic variability among the human and the avian strains of ST95. Thus, the genetic variability could be useful to separate these ST95 strains in the future in accordance with their host. This would be of utmost importance in terms of transmission events and risk assessment. On the other hand, these strains of ST95 may also have some host specific phenotypes that could differentiate these strains according to their host based on phenotypic characters.

Hence, the main aim of our study was to identify certain genetic markers discerning ST95 strains. We concentrated on single nucleotide polymorphism (SNP) based molecular genetic markers. The concept was to analyze both the Maximum common genome (MCG), which by definition means the set of genes that are shared by all the members of bacterial species and to also analyze the flexible or accessory gene pool. In this concern, different approaches were put forward to identify certain candidate genes from the MCG and flexible genome for which SNP typing could be performed to identify genetic markers. The second aim was to identify certain phenotypic characteristics in strains from ST95 complex that would segregate this group of strains according to their specific human/avian host respectively. Different phenotypic characteristics like adhesion and invasion capabilities, biofilm forming capabilities, expression of curli and cellulose via long term biofilm colony formation, haemagglutination assay and serum resistant phenotype of ST95 strains towards chicken and human serum were accessed, so as to investigate if this group of strains could be differentiated from each other based on their phenotypic behavior.

Also the genome flexibility among the avian ExPEC strains may trigger avian strains to tune themselves according to different hosts in particular human host. Thereby, widening their host spectrum for infection and crossing the species barrier and infecting other hosts and emerging as more successful pathogens.

2. Materials

2.1 Chemicals

Chemicals, reagents and solutions used during different experiments were procured from Roth GmbH & Co., Karlsruhe, Germany; BioRad, Munich, Germany; Invitrogen, Groningen, Germany; and Sigma, Deisenhofen, Germany.

Enzymes and oligonucleotide primers were purchased from Invitrogen GmbH, Karlsruhe-Germany, New England Biolabs GmbH, Frankfurt, Germany, Sigma-Aldrich Chemie GmbH, Munich-Germany and Thermo Fisher Scientific GmbH, Schwerte- Germany.

2.2. Materials for microbiological experiments

2.2.a. Bacterial strains

In total, 116 ExPEC strains previously typed as ST95 including 61 avian strains (58 APEC strains and 3 commensal strains) and 55 human strains (24 UPEC, 15 NMEC, 4 SePEC and 15 commensal strains) available at our institute were investigated in different experiments in this thesis. Additionally, 140 strains (isolated from dogs, cats, sheep and cattle) with different clinical implications were used for screening of *svg.I* gene (Bidet et al. 2007). These strains were isolated over a period of 2001-2010 mostly from Germany while others from USA, Finland, France, Canada, Jordan and The Netherlands.

Table 1: Strains from different hosts with various clinical outcomes screened for ST95-suspected gene sequence *svg.I*

Host	Disease	No. of Samples
Dog	Urinary tract infection, Prostatitis	50
Cat	Urinary tract infection	50
Cattle	Mastitis	30
Sheep	Diarrhea	10

All bacterial strains were cultured at 37 °C in Luria-Bertani (LB) broth and on LB agar plates with appropriate antibiotics when required in the following concentrations: Kanamycin (Kan), 50 µg/ml; Ampicillin (Amp), 50 µg/ml. Bacterial strains were stored at -80 °C in LB broth with 10 % (v/v) glycerol until further use. Initially, all glycerol stocks were streaked on sheep blood agar plate whenever used.

2.2.b. Growth media

The following growth media were used for the experiments performed in this thesis

Luria Bertani (LB) Broth

NaCl	5 g /L
Peptone (casein)	10 g /L
Yeast extract	5 g /L
pH	7.0

Autoclaved at 121°C, 15 min

Luria Bertani Agar

NaCl	5 g /L
Peptone (casein)	10 g /L
Yeast extract	5 g /L
Agar	15 g /L

Autoclaved at 121°C, 15 min

SOC media

Tryptone	20 g /L
Yeast extract	5 g /L
NaCl	0.5 g /L
KCl (20mM)	10 ml /L
NaOH (5N)	20 ml /L

Autoclaved at 121°C, 15 min

After cooling 5 ml of MgCl₂ (2 M) and 20 ml of filter sterilized Glucose (1 M) were added to the media.

M63 media

Ammonium Sulfate	2.0 g /L
Potassium Phosphate, Monobasic	13.6 g /L
Ferrous Sulfate x 7H ₂ O	0.5 mg /L

Autoclaved at 121°C, 15 min

Additional supplements were added after autoclaving:

Filter sterilized 20% solution of glycerol	10 ml /L
Sterile 1M MgSO ₄ solution (autoclaved)	1 ml /L

SOB media

Bacto Tryptone	20 g/ L
Bacto Yeast Extract	5 g/ L
5M NaCl	2 ml/ L
1M KCl	2.5 ml/ L
1M MgCl ₂	10 ml/ L
1M MgSO ₄	10 ml/ L

Autoclaved at 121°C, 15 min

Congo red Span Agar (with NaCl)

Tryptone	10 g/L
Yeast extract	5 g /L
NaCl	5 g/L
Span Agar	18 g/L

Autoclaved at 121°C, 15 min

Congo red Span Agar (without NaCl)

Tryptone	10 g/L
Yeast extract	5 g /L
Span Agar	18 g/L

Autoclaved at 121°C, 15 min

When the media (with and without NaCl) cooled down to 55°C, 20 ml/L Congo red solution (filter sterilized) was added to agar media and poured (40 ml) in Petri plates.

Congo red solution

Congo red	2 mg /ml
Coomassie-Brilliant-Blue G250	1 mg /ml

Dissolved in ddH₂O and filter sterilized.

2.3 Materials for molecular biology experiments**Solutions****Solutions used for agarose gel electrophoresis.****1.5% Agarose gel**

Agarose	1.5 g
1 x Tris Borate EDTA (TBE) Buffer	100 ml

Ethidium Bromide solution	2 μ l
---------------------------	-----------

0.8% Agarose gel

Agarose	0.8 g
1 x Tris Borate EDTA (TBE) Buffer	100 ml

Ethidium Bromide solution	2 μ l
---------------------------	-----------

0.4% Agarose gel

Agarose	0.4 g
1 x Tris Borate EDTA (TBE) Buffer	100 ml

Ethidium Bromide solution	2 μ l
---------------------------	-----------

10 x Tris Borate EDTA (TBE) Buffer

Tris base	108 g/L
-----------	---------

Boric acid	55 g/L
------------	--------

0.5M EDTA (pH 8.0)	40 ml
--------------------	-------

Autoclaved at 121°C, 15 min

Loading dye

Glycerol	5 ml (50%)
----------	------------

Bromophenol Blue	10 mg (0.1%)
------------------	--------------

0.5M EDTA solution	2 ml (0.1 M)
--------------------	--------------

Tris pH 8.0	0.1 ml (10 mM)
-------------	----------------

Distilled water	2.9 ml
-----------------	--------

Markers

100 bp and 1 kb DNA ladder

(Thermo Fisher Scientific GmbH Schwerte, Germany)

Solutions for competent cell preparation

1 M Calcium Chloride (CaCl_2) solution

$\text{CaCl}_2 \cdot \text{H}_2\text{O}$	129 g/L
--	---------

Adjust the volume to 1000 ml with deionized / Milli-Q water.

1 M Magnesium chloride (MgCl_2) solution

$\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$	203.30 g/L
---	------------

Adjust the volume to 1000 ml with deionized / Milli-Q water.

Magnesium Chloride and Calcium Chloride solution (80 mM MgCl_2 . 20 mM CaCl_2)

(¹Thermo Fisher Scientific GmbH Schwerte, Germany, ²TaKaRa Göttingen, Germany)

Long range PCR

High Fidelity PCR Enzyme Mix, 5 u / μ l¹

10 x High Fidelity PCR Buffer with 15 mM MgCl₂¹

dNTPs PCR nucleotide Mix: dATP, dCTP, dGTP, dTTP²

(¹Thermo Fisher Scientific GmbH Schwerte, Germany, ²TaKaRa Göttingen, Germany)

Reagents for ligation

5 x T4 DNA Ligase buffer¹

T4 DNA Ligase enzyme¹

Nuclease free water¹

(¹Thermo Fisher Scientific GmbH Schwerte, Germany)

Reagents for Restriction Digestion

10 x Fast Digest buffer¹

Fast Digest enzyme XhoI¹

Fast Digest enzyme SalI¹

(¹Thermo Fisher Scientific GmbH Schwerte, Germany)

Oligonucleotides

All oligonucleotide primers listed in the table below were procured from MWG Biotech, Ebensburg, Germany.

Table 2: Oligonucleotide primers used in this work

Primer number	Primer Name	Primer Sequence (5'-3')	Temperature	Reference
IMT4197	<i>svg.1</i> – For	TCCGGCTGATTACAAACCAAC	55°C	(Bidet et al. 2007)
IMT4198	<i>svg.2</i> – Rev	CTGCACGAGGTTGTAGTCCTG	55°C	(Bidet et al. 2007)
IMT5224	<i>fimH</i> – For	CACTCAGGGAACCATTTCAGGCA	55°C	This study
IMT5225	<i>fimH</i> – Rev	GCGTACCTGCATTAGCAAT	55°C	This study
IMT5916	<i>fimE</i> – For	GCCGGATTATGGGAAAGA	57°C	This study
IMT5917	<i>fimE</i> – Rev	GCCGCTGTAGAACTCAGGG	57°C	This study
IMT5918	<i>fim</i> operon – For	GGGGTCGACATGAAGAATAAG GCTGATAAC	60.4°C	This study
IMT5919	<i>fim</i> operon – Rev	GTGCTCGAGTTATTGATAAACA AAAGTCAC	60.4°C	This study

Primer number	Primer Name	Primer Sequence (5'-3')	Temperature	Reference
IMT6006	<i>ybjD</i> – For	AGCGCGTTGAAATTGTGGG	55°C	This study
IMT6007	<i>ybjD</i> – Rev	GCCCATCAGAGAGATTTTGCG	55°C	This study
IMT6008	<i>tatB</i> – For	GTAAAAACGGTAGCGGGCTG	57°C	This study
IMT6009	<i>tatB</i> – Rev	AGCCTGGCGTAATTCATCCA	57°C	This study
IMT6037	<i>hemH</i> – For	CAGACGCGTGGTTGATACCT	55°C	This study
IMT6038	<i>hemH</i> – Rev	GAATGCGTGCCAGTTCATCC	55°C	This study
IMT6039	<i>hisJ</i> – For	TCTTTAACAGACGGGCCACC	55°C	This study
IMT6040	<i>hisJ</i> – Rev	AGCGAAGAAGATTGACGCCA	55°C	This study
IMT6041	<i>hycD</i> – For	GGTTCAACCAACACGCCAAG	55°C	This study
IMT6042	<i>hycD</i> – Rev	GCTGCTCTCCGGTATTACCC	55°C	This study
IMT6043	<i>lon</i> – For	TGAGCCGGGTGTAAACGATC	55°C	This study
IMT6044	<i>lon</i> – Rev	GCCAGTTTCAGCGGCATATG	55°C	This study
IMT6045	<i>mazG</i> – For	GCCGTTTTACCTGTTGCCAG	55°C	This study
IMT6046	<i>mazG</i> – Rev	TTACCGGCTTTAATGCGTGC	55°C	This study
IMT6047	<i>rbn</i> – For	TTTTCGCCGCTTTTCCCATG	55°C	This study
IMT6048	<i>rbn</i> – Rev	CGACAACAGCAGCGGAAAAA	55°C	This study
IMT6049	<i>rnfA</i> – For	CGCCTGGCTTATCGATACGT	55°C	This study
IMT6050	<i>rnfA</i> – Rev	GCAATGGCATTACCGCGAAA	55°C	This study
IMT6051	<i>yehN</i> – For	TGTGAGCACTTTATCGGCTGA	55°C	This study
IMT6052	<i>yehN</i> – Rev	ATCGTGATCGTTGCCAATGG	55°C	This study
IMT6053	<i>yciM</i> – For	TGCTGAATTGATGCTTGCGG	55°C	This study
IMT6054	<i>yciM</i> – Rev	GCCATCAAGACCGCGAATTG	55°C	This study
IMT6055	<i>yniC</i> – For	AGGTTTATTGGTGGGGCTGG	55°C	This study
IMT6056	<i>yniC</i> – Rev	CGAGAAGGTCTTTTGCGGTG	55°C	This study

Table 3: List of kits

Master Pure™ Genomic DNA Purification Kit	Epicentre, Biozym, Hessisch Oldendorf, Germany
TOPO TA Cloning ® Kit	Invitrogen GmbH, Karlsruhe, Germany
QIAquick Gel extraction Kit	Qiagen GmbH, Hilden, Germany

Plasmid - pET 45b (+) vector plasmid (Novagen) was used for cloning experiments.

2.4. Materials for cell biology experiments

Reagents used in cell culture

10 X Phosphate Buffered Saline (PBS)

NaCl 80 g/L

KCl	2.0 g/L
Na ₂ HPO ₄	14.4 g/L
KH ₂ PO ₄	2.4 g/L

Dissolved in 800 ml distilled H₂O and pH adjusted to 7.4 and volume made up to 1 L.

Autoclaved at 121°C, 15 min

1 x PBS

Diluted from 10 x PBS in ddH₂O and then autoclaved at 121°C, 15 min.

0.1 % Triton X 100 solution

0.1 ml Triton X 100 dissolved in 99.9 ml 1 x PBS (sterile)

Fetal bovine serum (FBS)³

DMEM/ Ham's F-12 media (1:1) with stable glutamine³

Penicillin/ Streptomycin solution³

Gentamycin³

(10 x) Trypsin (1:250)/EDTA solution (0.5 %/0.2 %) ³

(³Biochrom AG, Berlin, Germany)

Other solutions

0.9 % NaCl solution + 5 % D- Mannose solution

2.5 Genomes for bioinformatic analysis

The following genomes of ExPEC strains (Table 4) were available at the start of this thesis and were used for the bioinformatic analysis to identify candidate genes that were used for SNP typing (Section 3.4.a and 3.4.b).

Table 4: Whole genome sequenced extraintestinal pathogenic *E. coli* (ExPEC) strains included for core genome comparison

<i>E. coli</i> strains	Host	Pathovar/ group	Serovar	Size (Mbp)	ST	STC	GenBank/ website
APEC_01	Chicken	APEC	O1:K1	5.1	95	95	CP000468
IHE3034	Human	NMEC	O18:K1:H7	4.1	95	95	CP001969.1
S88	Human	NMEC	O45:K1	5.2	95	95	CU928161
UTI 89	Human	UPEC	O18:K1:H7	5.1	95	95	CP000243
536	Human	UPEC	O6:K31:H15	4.9	127	127	CP000247
F11	Human	UPEC	O6:H31	5.2	127	127	AAJU00000000
RS218	Human	NMEC	O18:K1:H7	5.2	5.3	95	Genome.wisc.edu
IAI39	Human	UPEC	O7:K1	5.1	62	none	CU928164.2
UMN026	Human	UPEC	O17:K52:H18	5.4	597	69	CU928163.2
ABU83972	Human	ABU	Orough:K5:NM	5.1	73	73	CP001671
CFT073	Human	UPEC	O6:K2:H1	5.2	73	73	AE014075

APEC = Avian pathogenic *E. coli*; UPEC = Uropathogenic *E. coli*; NMEC = Neonatal meningitis *E. coli*; ABU = Asymptomatic Bacteriuria

2.6 Statistical analyses

Statistical analyses of the results of adhesion/invasion assays and biofilm formation assay (96-well microtiter plate method) were carried out using the nonparametric Mann-Whitney U test, which was implemented in the Statistical Package for the Social Sciences (SPSS, version 10.0) and Wilcoxon signed rank test using R software (R version 2.14.0). Statistical analyses for haemagglutination assay, serum assay was carried out using chi-square test using Statistical Package for the Social Sciences (SPSS, version 10.0). Statistical analyses for long term colony formation assay were carried out using multinomial logistic regression analysis using Statistical Package for the Social Sciences (SPSS, version 10.0).

Methods

3.1 Microbiology experiments

3.1.a Biofilm formation assay

3.1.a.i. 96-well microtiter plate method

This method was used to investigate biofilm forming capacities of ST95 strains isolated from human and avian hosts using a previously described method with some modifications (Wang et al. 2011, Hussain et al. 2012). W3110 (*E. coli* K-12 derivative) (Jackson et al. 2002, Colon-Gonzalez et al. 2004) was used as a positive control while AAEC189 (*E. coli* K-12 derivative) (Wang et al. 2011) served as a negative control. Bacterial inoculum was taken from frozen stock and streaked on blood agar plates and incubated overnight. Single colony from overnight incubated plate was picked and inoculated in 5 ml media (M63 and LB medium) and then incubated overnight at 37°C in a shaking incubator. All overnight cultures were diluted 1:10 in LB and M63 medium and the OD₆₀₀ was determined by a photometer. Cultures were diluted to an OD₆₀₀ = 0.05 in fresh M63 and LB medium. 200 µl of the diluted culture was pipetted in triplicates into 96-well microtiter plates, covered with a permeable film, and incubated for 24 and 48 hrs at 28°C without shaking. After 24 hrs of incubation, growth was determined by measuring the OD at 600 nm in an ELISA reader. After reading the plate, spent media was gently aspirated and the wells were washed three times with 300 µl of double distilled water and air-dried. 250 µl 99% methanol was then added to each well and incubated for 15 minutes to fix the attached bacteria. Methanol was aspirated and the plate was air-dried. Wells were then stained with 250 µl of 0.1% crystal violet solution for 30 minutes. Crystal violet was then removed and the plates were again washed three times with 300 µl of double distilled water and air-dried. 300 µl of ethanol:acetone (80:20) solution was added to the air-dried wells to re-solubilize the stain. The plate was placed on a horizontal shaker at 200 rpm until all the stain was solubilized (about 30-45 min). The crystal violet staining was measured by absorbance at OD₅₇₀ in the ELISA reader. The plate that was incubated for 48 hrs was also processed as mentioned above. Biofilm formation was calculated using the formula $SBF = (AB - CW) / G$, in which *SBF* is the specific biofilm formation, *AB* is the OD₅₇₀ nm of the attached and stained bacteria, *CW* is the OD₅₇₀ nm of the stained control wells containing only bacteria-free medium (to eliminate unspecific or abiotic OD values), and *G* is the OD₆₀₀ nm of cell growth in media.

The degree of biofilm production was classified in three categories: weak ($SBF \leq 0.5$), moderate ($0.5 > SBF \leq 1$), and strong ($SBF > 1$) (Margarita et al. 2009).

3.1.a.ii. Long term biofilm formation

A single colony was inoculated in 5 ml of LB media and incubated overnight at 37°C in a shaker incubator. 1.5 µl of the overnight culture was pipetted and a small single drop was dropped on Congo red plate prepared with and without salt. The drop was allowed to dry on the plate and later incubated at different temperatures of 28°C and 37°C. IMT26949 (UPEC strain isolated from dog with urinary tract infection) was used as a positive control to check the expression of curli and cellulose. The relationships among different growth conditions and colony morphotypes and strains isolated from human and avian host was determined using multinomial logistic regression analysis using Statistical Package for the Social Sciences (SPSS, version 10.0).

3.1.b Haemagglutination assay

Overnight culture was prepared by inoculation of a single colony in LB broth media and incubated overnight at 37°C in a shaker incubator. Chicken red blood erythrocytes (RBCs) were diluted in 0.9% NaCl solution with and without 5% D-Mannose. Two 30 µl drops of the bacterial culture were added onto the glass slide and 30 µl each, of diluted chicken RBC's (with and without D-Mannose) suspension was added separately to the bacterial culture drops on glass slide. Bacterial culture with chicken RBC's was mixed properly with sterile tooth picks and incubated at room temperature for 2 minutes. Visible clumps conferred agglutination in chicken RBC's, which was inhibited in presence of mannose. Chi-square test or χ^2 test was performed using Statistical Package for the Social Sciences (SPSS, version 10.0) to check the significant difference between strains showing haemagglutination.

3.1.c Serum assay

Serum assay was performed with chicken and human serum with some modification from Ewers et al. (2009). Fresh bacterial cultures were streaked from stock cultures on blood agar plates and incubated overnight at 37°C. A single colony from the overnight plate was inoculated in 5 ml LB media and incubated overnight at 37°C in a shaker incubator. Following day, 5 µl of the overnight culture was added to 495 µl of LB broth and incubated for 1.5 hours at 37°C in a shaker incubator. After incubation, the culture was centrifuged at 9,000 rpm for 3 minutes and

the supernatant was discarded while the pellet was resuspended in 1 x PBS. 270 µl of the normal serum was added to each well of 96-well plate in triplicate and 30 µl of the resuspended bacterial inoculum was added carefully to the serum in the 96-well plate (inoculation plate) and mixed properly. 270 µl of 1 x PBS was added to each well of the other 96-well plate to prepare dilutions (dilution plate). 30 µl of the serum and bacterial cell mixture from the inoculation plate was taken out and added to the 270 µl PBS in the dilution plate and further dilutions were made using multi-channel pipette. After proper dilutions 50 µl of all the dilutions were dropped in forms of small drops on to LB agar plates in duplicate and incubated overnight (0 hour count). The 96-well inoculation plate was incubated at 37°C in a shaker incubator for 4 hours at 100 rpm. After 4 hours of incubation 30 µl of the serum was again taken out and diluted as mentioned earlier (4 hour count). Next day, the number of colony forming units was determined for 0 hour and 4 hours. . Chi-square test or χ^2 test was performed using Statistical Package for the Social Sciences (SPSS, version 10.0) to check the significant difference between strains isolated from humans and birds showing serum sensitive and serum resistant phenotypes.

3.2 Molecular biology experiments

3.2.a Screening of available strains from different hosts for sequence type 95

Bacterial cell crude lysates preparation (Heat – lysis method) - A single colony of strains (Table 1) from an overnight incubated blood agar plate was resuspended in 200 µl deionized water (MilliQ water) in 0.5 ml centrifuge tube and incubated at 95°C for 10 minutes in a waterbath. After incubation the tubes were cooled down and centrifuged at 10,000 rpm for 10 minutes. The supernatant that contained genomic DNA was used as a template for PCR for identifying the presence of *svg.1* “specific for virulent subgroup (*svg*)” gene. This gene had previously been suggested to specifically identify strains belonging to ST95 (Bidet et al. 2007).

PCR reaction mixture

Deionized water (MQ H ₂ O)	12.75 µl
10 x PCR Reaction buffer	2.5 µl
Primer IMT1497 <i>svg.1</i> For (10 pmol)	0.25µl
Primer IMT1498 <i>svg.2</i> Rev (10 pmol)	0.25µl
MgCl ₂ (25mM)	2 µl
dNTP mix (10 mM)	0.2 µl
<i>Taq</i> polymerase (5U/ µl)	0.05 µl
DNA template	2 µl

PCR conditions

95°C	5 min	
94 °C	30 sec	30 cycles
55 °C	30 sec	
72 °C	30 sec	
72 °C	10 min	

PCR products were loaded on a 1.5% agarose gel and samples positive for *svg.1* gene were determined by the presence of a 434bp band.

3.2.b *fimH* gene amplification

PCR was performed for 116 ST95 strains available at our institute strain collection for SNP typing (see 3.4.d).

PCR reaction mixture

Deionized water (MQ H ₂ O)	24.55 µl
10 x PCR Reaction buffer	3.5 µl
Primer IMT5224 <i>fimH</i> For (10 pmol)	0.3 µl
Primer IMT5225 <i>fimH</i> Rev (10 pmol)	0.3 µl
MgCl ₂ (25mM)	2.0 µl
dNTP mix (10 mM)	0.2 µl
<i>Taq</i> polymerase (5U/ µl)	0.05 µl
DNA template	4.0 µl

PCR conditions

95°C	5 min	
94 °C	30 sec	30 cycles
55 °C	30 sec	
72 °C	90 sec	
72 °C	10 min	

4 µl of the PCR products were loaded on a 1.5 % agarose gel and samples positive for *fimH* gene were determined by the presence of a 1000 bp band and were later sequenced both in forward and reverse direction.

3.2.c Cloning of prominent *fimH* alleles

Some of the *fimH* alleles were more prominent among ST95 strains; hence, cloning of the *fim* operon from some strains of frequent *fimH* alleles was performed, so as to investigate if these predominant alleles provide some advantage to the bacteria.

3.2.c.i. Extraction of Genomic DNA

Bacterial genomic DNA was isolated from the overnight culture incubated at 37°C at 200 rpm in a shaker incubator, using 'MasterPure DNA Purification Kit for Blood Version II' (Epicentre Technologies) as recommended by the manufacturer. Once the pellet was dried, it was resuspended in 200 µl Millipore-Water and incubated for 1 hr at 65°C or overnight at room temperature and later quantified by Nano-Drop and was stored at -20°C until further use.

3.2.c.ii. Preparation of competent cells

Electro competent cells

One milliliter of the overnight culture was added to 100 ml of LB media or SOB media and incubated at 37°C in a shaker incubator till the OD₆₀₀ reached 0.6-0.7 nm. The culture was then kept on ice for 30 minutes and from here all the steps were performed on ice to maintain cold temperature. The culture was centrifuged at 4000 rpm for 20 minutes at 4°C using a 50 ml centrifuge tube. After centrifugation the supernatant was discarded carefully and the pellet was resuspended in 50 ml of ice cold ddH₂O and centrifuged as mentioned earlier. After centrifugation the supernatant was discarded and the pellet was again resuspended in 25 ml of ice cold ddH₂O and centrifuged as mentioned earlier. Again the supernatant was poured off and resuspended in 10% glycerol in ddH₂O (2 ml for each 50 ml of bacterial culture) and centrifuged at 4000 rpm for 20 minutes at 4°C. Finally, the supernatant was removed and the pellet was resuspended in 10% glycerol in ddH₂O (1 ml for each 50 ml of bacterial culture). 100 µl of this bacterial cell suspension was aliquoted in pre chilled 1.5ml eppendorf tubes and were snap freezed in liquid N₂ and then the tubes were stored in -80°C.

CaCl₂ competent cells

A single bacterial colony was inoculated in 20 ml LB and incubated at 37°C in a shaker incubator at 200 rpm to an OD₆₀₀ of 0.4-0.6 and then inoculated in 100 ml LB media and incubated till the OD₆₀₀ reached 0.4 -0.6 and then the culture flasks were placed on ice for 30 minutes. The bacterial culture was centrifuged in 50 ml centrifuge tube for 15 minutes at 4000 rpm at 4°C. After centrifugation the supernatant was discarded and the bacterial pellet was resuspended in 30 ml of ice cold MgCl₂-CaCl₂ solution (80 mM MgCl₂. 20 mM CaCl₂) by swirling or gentle vortexing and again centrifuged as mentioned above. The supernatant was again discarded and the pellet was resuspended in 2 ml of ice cold 0.1 M CaCl₂ solution (containing 20% glycerol) for each 50 ml of culture. 100 µl of this bacterial cell suspension was

aliquoted in pre chilled 1.5 ml eppendorf tubes and were snap freezed in liquid N₂ and then the tubes were stored at -80°C.

3.2.c.iii. Long range PCR for amplification of DNA insert

This PCR was performed to amplify the DNA fragment (*fim* operon - CP000243.1|:4905700-4914457 *Escherichia coli* UTI89 ~8.7 kb- used as a reference for designing primers) to be used as an insert fragment in cloning. Strains form predominant *fimH* alleles (designated as I, II, III till XIV as determined by SNP typing of *fimH* gene Table 5) were used for cloning.

Table 5: Strains representing predominant *fimH* alleles that were used for cloning

No.	Allele No.	Strain
1.	I	IMT23830 (APEC_O1)
2.	II	IMT9241
3.	III	IMT12024
4.	IV	IMT8894
5.	V	IMT9264 (IHE3034)

PCR reaction mixture

Deionized water (MQ H ₂ O)	30 µl
10X High Fidelity PCR Reaction buffer	5 µl
Primer IMT5918 <i>fim</i> operon For (10 pmol)	5 µl
Primer IMT5919 <i>fim</i> operon Rev (10 pmol)	5 µl
dNTP mix (10 mM)	0.5 µl
High Fidelity Taq Polymerase	0.5 µl
DNA template (50 ng/µl)	4.0 µl
Total	50 µl

PCR Program

94°C	3 min	
94 °C	30 sec	10 cycles
60.4 °C	30 sec	
68 °C	8 min	
94 °C	30 sec	20 cycles
60.4 °C	30 sec	
68 °C	12 min	
68 °C	10 min	

The PCR products were loaded on a 0.4 % agarose gel and samples positive for *fim* operon were determined by the presence of 8.5 kb band.

3.2.c.iv. Ethanol precipitation of DNA

1/10 volume of 3 M sodium acetate and 10 volume of 100% ethanol was added to the DNA sample (product of long range PCR), mixed properly and stored at -20°C for at least 30 minutes. After that the sample was centrifuged at 12,000 rpm for 15 minutes. The supernatant was discarded carefully and the pellet was washed twice with 70% ethanol and then air-dried to remove ethanol traces and later resuspended in the appropriate volume of TE or water.

3.2.c.v. Isolation of plasmid by alkaline lysis method

A single colony was inoculated in 5 ml of LB media and incubated overnight in a shaker incubator at 37°C. 3 ml of the overnight culture was centrifuged in a 1.5 ml centrifuge tube for two minutes at 6,000 rpm and the supernatant was discarded. The pellet was resuspended in 200 µl of solution I containing T₂₅E₁₀ + 4 µl RNase (Tris-pH 7.5 to 8, 25 mM, EDTA 10 mM) by vortexing. 200 µl of solution II (0.2 M NaOH, 1% SDS) was then added to the tube and incubated at room temperature for 15 minutes. After incubation, 200 µl of solution III (3 M Potassium Acetate pH 5) was added and mixed by inverting the tube 10-12 times and incubated on ice for 20 minutes. The tube was then centrifuged at 12,000 rpm for 15 minutes. Clear supernatant was collected in a new tube and 1 ml of 100% ethanol was added and then the tube was centrifuged at 12,000 rpm for 15 minutes. The supernatant was discarded and the pellet was washed twice with 500 µl of 70% ethanol and air-dried completely to remove ethanol. The pellet was then resuspended in MQ water. The plasmids isolated were loaded on a 0.8 % agarose gel and samples positive for pET-45b(+) vector were determined by the presence of a 5.2 kb band.

3.2.c.vi. Restriction digestion of plasmid vector pET-45b(+) and DNA insert

Restriction digestion of the plasmid vector and DNA insert was performed using restriction enzymes as recommended by the enzyme manufacturers.

Fast digest buffer	1 µl
Plasmid vector	3 µl
XhoI enzyme	0.5 µl
SalI enzyme	0.5 µl
MilliQ water	5 µl

The reaction mixture was incubated at 37°C in a water bath and later the enzymes were inactivated by heating the reaction mixture at 60°C for 5 minutes.

3.2.c.vii. Purification of digested DNA and plasmid vector

Digested DNA was purified by ethanol precipitation as described earlier. Plasmid was purified using 'QIAquick Gel Extraction Kit protocol' (QIAGEN GmbH, Hilden, Germany) as recommended. Briefly, the digested plasmid was excised from gel and weighed in 2 ml eppendorf tube. Three volumes of QG buffer were added to 1 volume of gel and were incubated at 50°C till the gel was dissolved completely. Then, 1 volume of iso-propanol was added to the tube and mixed by inverting the tube. The mixture was added to the spin column (provided in kit) and centrifuged at 13,000 rpm for 1 minute. 0.75ml of buffer PE was added to the column for washing and then centrifuged as mentioned earlier. The spin column was then placed in a 1.5 ml centrifuge tube and EB buffer or MilliQ water was added to the column and allowed to stand at room temperature for 2 minutes and then centrifuged at 13,000 rpm for 1 minute to elute the plasmid. Concentrations were determined using Nanodrop (NanoDrop 1000 Spectrophotometer from Thermo Scientific, GmbH Schwerte, Germany).

3.2.c.viii. Ligation of plasmid vector and DNA insert

The ligation reaction was set up according to the manufacturers recommendations, considering a vector insert ratio of 1:10.

T4 Ligase buffer	1 µl
Plasmid vector (pET-45 b(+))	1 µl
DNA Insert	5 µl
T4 DNA Ligase enzyme	1 µl
MilliQ water	2 µl
Total	10 µl

The ligation mixture was incubated at room temperature for 1 hour and then overnight at 4°C and later used for transformation into competent cells.

3.2.c.ix. Transformation of competent cells

Heat shock method

An aliquot of CaCl₂ competent cells was thawed on ice and 2-5 µl of the ligation mixture was added to the thawed cells and incubated on ice for 20 minutes. Then the bacterial cells were given heat shock by placing the tube in a waterbath at 42°C for 90 seconds and then left on ice for 1-2 minutes. 0.8 ml of LB media was added to the tube and the tube was placed in a shaker

incubator at 37°C for 45 minutes. After incubation, the tube was centrifuged to spin down the cells and excess of media was discarded. 250 µl media was left in the tube and the pellet was resuspended in these 250 µl LB media and then spread on LB plates having appropriate antibiotic and incubated at 37°C overnight (12-16 hrs). After overnight incubation the colonies were screened for clones carrying the DNA insert by colony PCR.

Electroporation method

An aliquot of electro-competent cells from -80°C was thawed on ice and 2-5 µl of the ligation mixture was added to thawed competent cells. The cells along with ligation mixture were added to a pre-cooled cuvette and placed in the electroporator and electroporation was carried out at 2,500 V/sec. After electroporation 200 µl of pre-warmed LB media (37°C) was added to the cuvette and then the cell suspension added to 1 ml LB media in 1.5 eppendorf tube and incubated for 3 hrs at 37°C in a shaker incubator. After incubation, the tube was centrifuged at 5,000 rpm for 2 minutes. The supernatant was discarded and the pellet was resuspended in 200 µl of LB media. 100 µl of this bacterial suspension was spread on LB agar plate with appropriate antibiotic and incubated overnight at 37°C. To the other 100 µl of bacterial suspension 900 µl of LB media was added and incubated at 37°C in a shaker incubator to plate on next day if required. After overnight incubation the colonies were screened for clones carrying the DNA insert by colony PCR.

3.2.c.x. Screening of colonies for DNA inserts.

A master-mix was prepared with specific primers for the DNA insert (specific primers for *fimE* gene which is regularly present in the *fim* operon) and aliquoted in cold PCR tubes kept on ice. To each tube containing the PCR reaction, a small amount of colony was added with a tip and pipetted up and down to mix. Sufficient mixing resulted in complete cell lysis and high yields.

PCR reaction mixture

Deionized water (MQ H ₂ O)	21.75 µl
10 x PCR Green Reaction buffer	3 µl
Primer IMT5916 <i>fimE</i> For (10 pmol)	0.4 µl
Primer IMT5917 <i>fimE</i> Rev (10 pmol)	0.4 µl
dNTP mix (10 mM)	0.4 µl
DreamTaq Green DNA Polymerase (5u /µl)	0.05 µl

PCR conditions

95°C 5 min
94 °C 30 sec
57 °C 30 sec 30 cycles
72 °C 45 sec
72 °C 10 min

The PCR products were loaded on a 1.5 % agarose gel and samples positive for *fimE* gene were determined by the presence of a 601 bp-band. Colonies found positive for PCR were inoculated in 5 ml LB media for plasmid isolation which were digested for clone conformation.

3.2.c.xi. Agarose gel electrophoresis

Agarose gel electrophoresis was used to determine the presence and specific size of PCR products. Agarose was weighed depending on the percentage of gel to be prepared (e.g. for 1.5% gel add 1.5 gram agarose in 100 ml of TBE buffer) and added to 1 x TBE buffer and melted in microwave till the solution becomes clear. Once the agarose was cooled to 50-55°C, EtBr solution was added and the agarose was poured in the gel tray with sealed ends and comb placed properly. When the gel was solidified, loading dye was added to the PCR tubes and mixed with the samples and each sample was loaded in a separate well of the gel, with proper marker in one well. Electrodes were connected correctly and the run was applied according to the gel percentage and size of gel (e.g. for 0.8 to 1.5 % agarose gel 100 to 130 volts).

3.2.d TOPO cloning

TOPO cloning was performed using ‘TOPO TA Cloning Kit’ (Invitrogen Life Technologies) as recommended by the manufacturers. Taq polymerase has a non template dependent terminal transferase activity that adds a single deoxyadenosine (A) to the 3’ ends of PCR products. The linearized vector supplied in the kit has single, overhanging 3’ deoxythymidine (T) residues. This facilitates PCR inserts to ligate efficiently with the vector. 50 µl PCR reaction was setup for this procedure using the cycling parameters according to the primers with an extra 30 minutes extension at 72°C after the last cycle to ensure that all PCR products were full length and 3’ acetylated. Once the desired PCR product was obtained (product checked on agarose gel) ligation reaction was setup with linear vector according to manufacturer’s instruction.

Ligation mixture

Fresh PCR product	4 μ l
Diluted salt solution	1 μ l
MilliQ water	0.75 μ l
TOPO vector	0.25 μ l
Total	6 μ l

The ligation mixture was kept at room temperature for 1 hour and then transformed in electro-competent or chemical competent TOP10 *E. coli*.

3.3 Cell culture experiments

3.3.a. Adhesion and invasion assay

3.3.a.i) Reviving of cell line from liquid nitrogen.

The vial from liquid nitrogen freezer was removed and thawed by gentle agitation in a 37°C water bath until ice crystals were melted and then the contents of the vial were transferred into a sterile 1.5 centrifuge tube and 0.5 ml of complete media (DMEM F-12 media with 10% FCS and 1% penicillin and streptomycin antibiotic mixture for MDCK-1 (*Madin-Darby Canine Kidney-1*) and HEK-293T (Human embryonic kidney-293T) cells) was added to the tube and centrifuged at 1000 rpm for 10 minutes. Culture vessel (T-25 flask) was prepared containing 4 ml of the appropriate culture medium at proper temperature. After centrifugation supernatant was discarded and the pellet was resuspended gently in 1 ml of complete media to avoid cell clumps. The cell suspension was transferred into the culture vessel and was swirled gently in direction of making 8. The culture vessel was then incubated in 5% CO₂ incubator at 37°C and was examined after every 24 hrs to check if subculture was required.

3.3.a.ii) Subculturing of adherent cells

Trypsin-EDTA solution, complete growth medium, 1 x PBS and antibiotic solution were maintained at 37°C. The cell culture medium was removed from the flask without disturbing monolayer and the monolayer was washed once with sterile 2 ml 1 x PBS. Two to 3 ml of Trypsin-EDTA solution was added to the monolayer and incubated at 37°C for appropriate time till the cells got detached. Six to 8 ml of complete growth medium was then added to the cell suspension to inactivate the trypsin. The cell suspension was collected in a centrifuge tube and centrifuged at 1,000 rpm for 10 minutes. The supernatant was discarded and the pellet was resuspended in appropriate volume of media. 4 ml of complete media was added to a new T-25 flask and cell suspension was added to the flask and placed back into the incubator and

examined next day to ensure that the cells have reattached and were growing actively. The medium was changed when the media color changed yellowish from pink.

Animal cell line (MDCK-1) and human cell line (HEK-293T) were used for adhesion and invasion assays. Monolayer of MDCK-1 and HEK-293T cells was maintained at 37°C in 5% CO₂:95% air (v/v) using Dulbecco's Modified Eagle's Medium supplemented with 10% fetal calf serum and 1% Penicillin and streptomycin antibiotic.

3.3.a.iii) Seeding the cells

To seed the cells for performing adhesion and invasion assays, a confluent (70–90 %) cell culture flask was taken and the entire medium was removed. The cell monolayer was washed once with sterile 1 x PBS. Trypsin- EDTA solution was then added to the monolayer and the flask was incubated at 37°C for 10 minutes. Six to 8 ml of complete growth medium was then added into the flask to the cell suspension to inactivate the trypsin. The detached cells were transferred to a 15 ml centrifuge tube and centrifuged at 1,000 rpm for 10 minutes. The supernatant was discarded and the cell pellet was resuspended in 5 ml fresh warm medium. Around 30-50µl of this suspension was added to each well of a 12-well cell culture plate containing 1 ml medium and incubated at 37 °C in 5 % CO₂ until a confluent monolayer was formed for performing assay.

3.3.a.iv) Performing adhesion and invasion assay

To quantify the association of bacteria with eukaryotic cells, adhesion assays were performed with some modifications from (Dogan et al. 2006). Briefly, bacteria were pelleted and resuspended in DMEM. Confluent monolayer of MDCK-1 and HEK-293T cells ($\sim 2 \times 10^7$) grown in 12-well plates were infected with the bacteria resuspended in DMEM at an MOI (multiplicity of infection) of 10 and incubated for 3 hours in a CO₂ incubator. After 3 hours of incubation at 37°C with 5% CO₂, the monolayer was washed three times with 1 x PBS and lysed with 1 ml of 0.1% Triton X-100 (in 1 x PBS) for 10 minutes. Upon lysis, lysates were collected and serially diluted and then plated on LB-agar plates and incubated overnight. The colonies were enumerated after overnight incubation. Adhesion was expressed as the total number of colony-forming units (CFU)/ml recovered per well. Each assay was repeated three times with three different biological samples. *E. coli* strain CFT073 strain was used as positive control and AAEC189 (a *fim*-negative *E. coli* K-12 derivative) as a negative control.

Bacterial invasive ability was assessed by performing invasion assays in MDCK-1 and HEK-293T cells with some modifications from Dogan et al. (2006) (Dogan et al. 2006). Confluent monolayer of MDCK-1 and HEK293T cells were grown in 12-well plates ($\sim 2 \times 10^7$) and infected with bacteria at an MOI of 10 as described in the adhesion assay. After 3 hrs of incubation, media was removed and fresh 1 ml of DMEM media (without serum) containing 50 mg/ml gentamicin was added to the monolayer and incubated for another 1.5 hours to kill extracellular bacteria. After incubation, the monolayer was washed three times with 1 x PBS and lysed with 1 ml of 0.1% Triton X-100 (in 1 x PBS) for 10 minutes. The lysates were collected and serially diluted and then plated on LB-agar plates. After overnight incubation the colonies were counted. Invasion was expressed as the total number of CFU/ml recovered per well. Each assay was repeated three times with three different biological samples. Avian *E. coli* strain IMT9294 (MT78) (Matter et al. 2011) was used as positive control and AAEC189 (*fim*-negative *E. coli* K-12 derivative) as a negative control.

3.4 Bioinformatics analysis

3.4.a Comparison of maximum common genome (MCG) sequences of 46 different *E. coli* strains

The strategy for distinguishing ST95 strains from different hosts initially aims at determining target candidate genes from maximum common genome that are presumed to be highly conserved among the bacterial species but might have established minor host-adaptive changes among certain subgroups during bacterial evolution. These genes could be used to identify single nucleotide polymorphisms (SNPs) discerning ST95 strains with respect to their host. At the beginning of my thesis work, only one avian ExPEC strain (APEC O1) was fully sequenced (Johnson et al. 2007), while a number of available ExPEC genomes were from human isolates (Table 4). For the comparison of maximum common genome among different *E. coli* strains, all four available ST95 ExPEC genomes were included, while in addition fully sequenced ExPEC strains belonging to other STs (all from human) as well as strains from other *E. coli* pathotypes and K-12 laboratory strains (follow Multiple alignment tree [Fig 4] for GenBank Accession numbers) were included to increase the robustness and reliability of the method.

As mentioned 46 fully annotated *E. coli* genomes were used for maximum common genome comparison to identify candidate genes suitable for SNP analyses. Coding sequences of these genomes were procured from NCBI (<http://www.ncbi.nlm.nih.gov/genome>) and were clustered using CD-HIT program (developed by Dr. Weizhong Li at Dr. Adam Godzik's Lab at the Burnham Institute (Sanford-Burnham Medical Research Institute). The filtering of the gene clusters that were related to the maximum common genome was then performed and these gene clusters were subjected to alignment using MUSCLE (Edgar 2004). Later, trees were generated using Maximum Likelihood method (Cam 1990). Therefore, genes that were conserved within the maximum common genome of different strains were identified.

Further analysis of the genes that were identified above was done for the four strains belonging to ST95 that are listed in Table 4, using Treeview program (TreeviewX Version 0.5.0) (Page 1996). This analysis was based on sequence variations among the genes within four strains of ST95 that designate the strains on different branches of the multiple alignment tree.

In this way, all the genes that showed variations in the sequences of the genes in avian ExPEC strain APEC O1 and the other three human ExPEC strains of ST95 were selected.

3.4.b *In silico* analysis of 46 genomes of *E. coli* strains including STC95 strains, based on recombination and mutation events to identify candidate genes for SNP analysis discerning ST95 strains according to specific human/avian host.

Another approach was also applied for the identification of genes for SNP analysis under the supervision of Dr. Haritha and was performed by Christoph Hartmann, who were members of the core bioinformatic facility of our institute. This analysis aimed at identifying recombination versus mutation events among the previously selected genes from maximum common genome shared by all the 46 *E. coli* genomes, in order to minimize the number of candidate genes included in subsequent SNP typing. The basic concept behind this idea is that genes undergoing frequent recombination cannot serve as good genetic markers. The above mentioned 46 *E. coli* genomes and 2169 gene clusters were used for this analysis and multiple sequence alignment of orthologous genes was performed using the ClustalW (ClustalW version 2.0) (Larkin et al. 2007) program with default settings. To construct a phylogenetic compatibility matrix, all orthologous gene sets were first concatenated into a single alignment. The evolutionary genomic approaches such as phylogenetic and substitution pattern methods were used to detect recombination events by examining the phylogenetic congruence among genes and also to

investigate the impact of recombination on the core genome of *E. coli*. We applied a method which tests for the rejection of a set of topologies, by a set of orthologous genes using the AU (Approximately unbiased) test (Shimodaira 2002, Susko et al. 2006). When a gene rejects a tree that is supported by the majority of the other genes, this gene is considered to be laterally transferred. We applied this approach to our data sets, testing topologies of the individual gene trees obtained by using Phylogenetic estimation using Maximum Likelihood (phyML version 3.0) (Guindon et al. 2010) and comparing it with the additional parsimony tree obtained with PHYLIP (Phylogeny Inference Package) (available online- provided by Joe Felsenstein, Department of Genome Sciences and the Department of Biology at the University of Washington) that was reconstructed from the concatenation of all genes. The site likelihood of each tree was then computed by the program Baseml (PAML package) (Yang 1997). The AU test was then applied using Consel (Shimodaira 2001) that is mainly designed for the selection of phylogenetic trees. Consel reads the output of phylogenetic packages like PAML, Phyml and calculates the *p*-value using several testing procedures like bootstrap probability and also calculates *p*-value based on AU test using multi scale bootstrap technique. Further identification of potential recombination evidence from each aligned orthologous group (at a nucleotide level) was performed using the MaxChi (Maximum Chi square test), Pairwise Homoplasy index (PHI) which has been indicated to be powerful in detecting recombination events (Smith 1992). Therefore, genes that undergo mutations and not recombination were selected.

3.4.c Identification of genetic markers by analyzing genes of core genome from nine fully sequenced ST95 strains genome.

In another approach, for identification of candidate genes for SNP typing, we analysed 1300 genes from the maximum common genome of nine strains that belong to ST95 complex (Table 6). Among these nine strains, the sequences of four ST95 strains (APEC_O1, S88, IHE3034 and UTI89) were available at NCBI and the other additional five strains (two strains [IMT11979, IMT13832] isolated from human, two strains isolated from birds [IMT9243, IMT15146], and one strain of canine origin [IMT20122]) that belong to ST95 were sequenced by our institute.

Table 6: *E. coli* ST95 complex strains used for genome comparison

Strain	Pathotype*	Host	Year	Region
APEC_O1	APEC	Bird	unknown	USA
IHE3034	NMEC	Human	1976	Finland
S88	NMEC	Human	unknown	France
UTI89	UPEC	Human	unknown	unknown
IMT13832	UPEC	Human	2001	Germany
IMT15146	Commensal	Bird	2008	Germany
IMT11979	UPEC	Human	2005	Germany
IMT20122	UPEC	Dog	2009	Netherlands
IMT9243	APEC	Bird	2005	Germany

*APEC = Avian pathogenic *E. coli*; UPEC = Uropathogenic *E. coli*; NMEC = Neonatal Meningitis *E. coli*

Ridom® SeqSphere Software (Ridom Bioinformatics, Ridom® GmbH, Munster, Germany) (Junemann et al. 2013) was used to analyse 1300 genes using default settings, in order to identify certain candidate genes that could separate the strains isolated from birds from that of the strains isolated from humans and a dog. Following this approach, certain genes were identified that differentiated strains of avian origin from strains isolated from humans and dog.

The genes that were identified earlier as candidate genes for SNP analysis were amplified using the below mentioned PCR conditions from the genomic DNA of 24 strains isolated from humans and 24 strains of avian origin belonging to ST95 from the strain collection of our institute and were sequenced and later analysed for specific SNPs.

PCR reaction mixture

Deionized water (MQ H ₂ O)	27.05 µl
10 x PCR Green Reaction buffer	4 µl
Primer IMT6006 – <i>ybjD</i> (10 pmol)	0.3 µl
Primer IMT6007 – <i>ybjD</i> (10 pmol)	0.3 µl
dNTP mix (10 mM)	0.3 µl
DreamTaq Green DNA Polymerase (5u /µl)	0.05 µl

PCR conditions

95°C	5 min	
94 °C	60 sec	30 cycles
55 °C	30 sec	
72 °C	45 sec	
72 °C	10 min	

The above reaction mixture was the same for all genes that were amplified with their respective primers; also identical PCR conditions were used for all the genes. Four microliter of the PCR products were loaded on a 1.5 % agarose gel and samples positive were later sequenced. The sequences were then analyzed by MEGA software (Molecular Evolutionary Genetics Analysis Version 5.05) (Tamura et al. 2011).

3.4.d SNP analysis of *E. coli* fimbrial adhesin gene

As the bioinformatics approach to identify suitable genes for SNP analysis from MCG was carried out, SNP analysis was also performed for *fimH* gene that has already been suggested suitable for discerning ExPEC groups (Tartof et al. 2007). FimH is part of a chromosomal operon and its gene product (adhesin) is responsible for the mannose-specific or receptor-specific binding of type I fimbriae.

PCR was performed for all ST95 strains available at our institute strain collection (see 3.2.a). The PCR products positive for *fimH* gene were sequenced both in forward and reverse direction. The sequences were then analyzed by MEGA software (Molecular Evolutionary Genetics Analysis Version 5.05) (Tamura et al. 2011).

4. Results

According to the publicly available database (<http://mlst.ucc.ie/mlst/mlst/dbs/Ecoli/> as observed in April 2013) sequence type 95 (ST95) contains strains, mostly isolated from humans and avian species with an exception of one canine strain. Therefore this ST could be designated as “Human-Avian” ST based on previous results from our group, where a large set of strains from cattle, swine, dogs and cats were mostly determined as non-ST95 strains, substantiating the limited host range of ST95 strains (Ewers et al unpublished data). This designation envisions that these strains are some kind of host limited. Efforts have been made to discern this group of ExPEC strains with respect to their clinical implications and host origin, but a precise definition and methods to study the risk assessment of these strains, in particular with regard to their zoonotic potential, has not been found today.

Currently it is not known whether all of these strains are zoonotic, or whether some strains are specific to one particular host, namely the human or an avian host. We hypothesized that such a host-specificity is reflected by a genetic variability among the human and the avian strains of ST95. If so, this genetic difference would be useful to discern these strains in the future in accordance with their host. This would be of utmost importance in terms of transmission events and risk assessment. On the other hand, these strains of ST95 may also have some host specific phenotypes that could differentiate these strains according to their host based on phenotypic characters. Hence, the main aim of our study was to identify certain genetic markers which would allow a rapid and specific risk assessment in the future. We concentrated on single nucleotide polymorphisms (SNP) that discern ST95 strains in a host specific manner. The concept was to analyze both the Maximum common genome (MCG) and the flexible or accessory gene pool.

Different approaches were put forward to identify certain candidate genes from the MCG and flexible genome for which SNP typing could be performed to identify genetic marker. The second aim was to identify certain phenotypic characteristics in strains from ST95 complex that would segregate this group of strains according to their origin from specific human/avian hosts. Different phenotypic characteristics like adhesion and invasion capabilities, biofilm forming capabilities, expression of curli and cellulose via long term biofilm colony formation, haemagglutination assay and serum resistant phenotype of ST95 strains towards chicken and

human serum, were accessed, so as to investigate if this group of strains could be differentiated from each other based on their phenotypic behavior.

For the identification of candidate genes that could be analyzed for SNP typing, initially the MCG of publicly available 46 *E. coli* genomes including strains of ST95 were procured from NCBI. In another approach the MCG of only those strains that belonged to ST95 and are available at NCBI were used along with additional five ST95 strains (two strains isolated from humans, two strains isolated from bird and one strain of dog) whole genome sequenced by our institute in the last year.

For the identification of specific phenotypes among the strains of ST95, 116 strains including 61 strains isolated from birds and 55 strains isolated from humans (Habilitation-Ewers 2012) were investigated to determine their phenotypic characters by performing biofilm formation assay, long-term biofilm formation assay, haemagglutination assay and serum sensitivity assay. The adhesion and the invasion capabilities of thirteen strains isolated from humans (six NMEC, six UPEC strains and one commensal strain) and eight strains from birds (six APEC strains and two commensal strains) out of 116 ST95 strains were also investigated. The strain selection for adhesion and invasion assay was based on their similar virulence gene pattern and clonal nature as analyzed by their ST, macro-restriction pattern of their genomes (Pulsed-field gel electrophoresis) that had been done before the start of this thesis (Habilitation-Ewers 2012). Thus, different phenotypic characters were investigated for the strains of ST95 available at our institute.

4.1. Screening of available strains from different animal hosts for sequence type 95

At the start of this thesis, a collection 120 strains of ST95 were available for our analyses. To include as many as possible ST95 strains, during the time of this thesis additional 140 strains from various hosts were screened for ST95 by PCR for a 434-bp fragment of the gene termed “specific for virulent subgroup (*svg*)” that had previously been suggested to specifically identify strains belonging to ST95 (Bidet et al. 2007). Most of these strains (Table 1) turned out to be negative for *svg* with the only exception of one strain isolated from a cat. However, subsequent MLST analysis assigned this strain to ST568 and not as expected to ST95. Thus, the published PCR is not specific for ST95. However, as ST568, which shares two alleles with ST95, belongs to the ST complex 95, this PCR seems to be useful as a pre-screen for strains of STC95.

4.2. Identification of genetic markers -- single nucleotide polymorphisms (SNPs) - that discern ST95 strains with respect to their host origin.

4.2.a. Comparison of maximum common genome (MCG) sequences of 46 different *E. coli* strains

A first step in the putative identification of SNPs in ST95 strains being specifically associated with specific hosts was the analysis of a maximum common genome (MCG). The MCG was initially identified from the whole annotated genomes of 46 *E. coli* strains, resulting in a total of 2,169 conserved genes that were shared by these 46 different *E. coli* strains. Multiple sequence alignment trees were obtained for all these genes. Using the program TreeviewX each of the single 2,169 gene trees were again analyzed to determine sequence variations in the sequences among four ST95 strains as listed in Table 7.

Table 7: Publicly available genomes of ST95 *E. coli* strains used for SNP-analysis

Strain	Pathotype*	Host	Accession number	Reference citation
APEC_O1	APEC	Bird	NC_008563.1	(Johnson et al. 2007)
IHE3034	NMEC	Human	CP_001969.1	(Moriel et al. 2010)
S88	NMEC	Human	NC_011742.1	(Touchon et al. 2009)
UTI89	UPEC	Human	NC_007946.1	(Chen et al. 2006)

ST95 strains having no sequence variability appeared to be on the same branch of multiple sequence alignment trees for a given gene cluster while differences in the sequences of a gene, separated the four ST95 strains on different branches of the multiple sequence alignment tree. Examples for two genes are given in Figure 4.

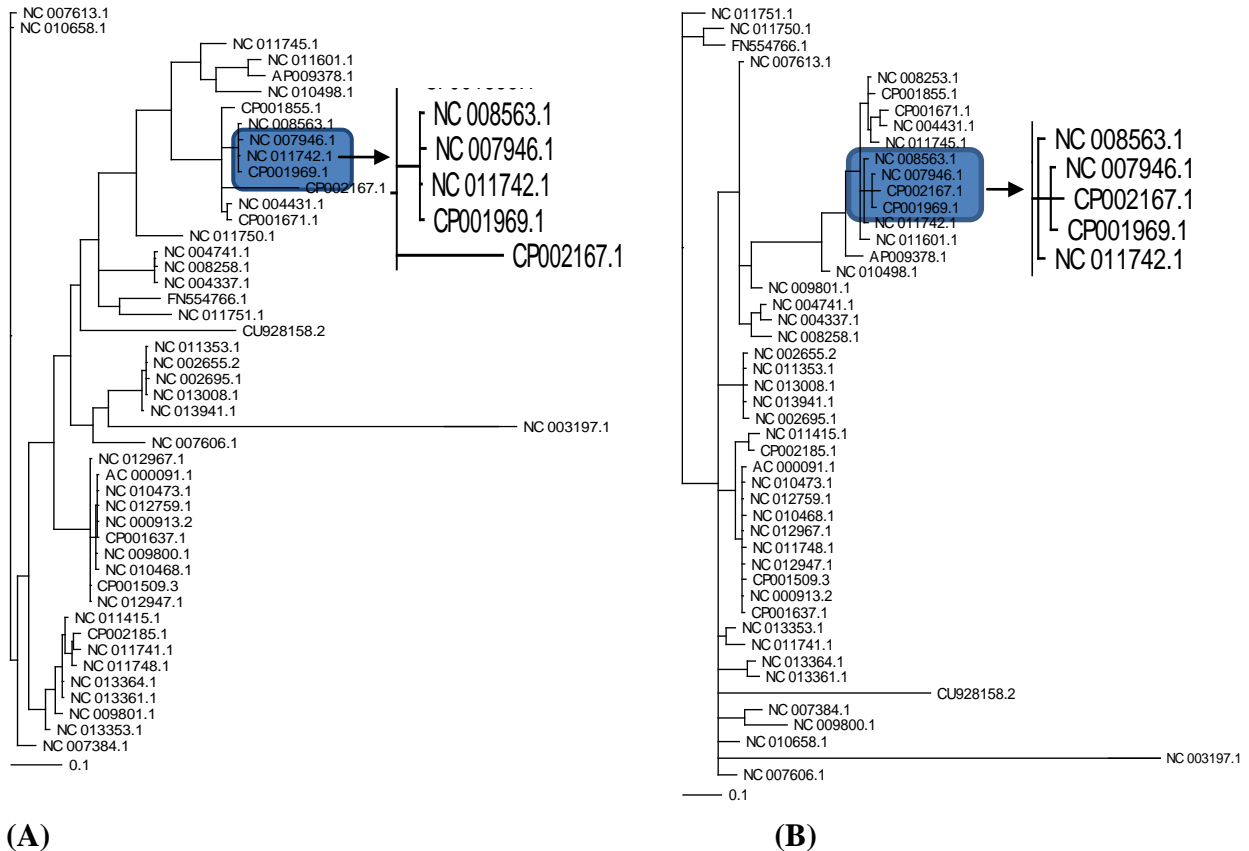


Figure 4: Dendrograms exemplarily depicting multiple alignment trees based on the sequences of two core genome genes of the 46 *E. coli* strains, the genomes of which were annotated and publicly available at the start of the thesis. **(A)** Gene cluster glutamate synthase represents no differences in the sequence and the four strains appear on the same branch. **(B)** Gene cluster cysteine sulfinatase desulfonase represents strains on different branches due to differences in the sequences among the strains.

This analysis of 2,169 genes revealed 172 genes that showed differences in the sequences of the genes of avian ExPEC strain APEC O1, designating this strain on a different branch when compared to the other three human ExPEC strains. Similarly, strain IHE3034 (NMEC), isolated from a human case of meningitis, possessed 73 genes that showed sequence variations when compared to the other ST95 strains. S88 (NMEC) had 93 genes and strain UTI89 (UPEC), isolated from a human case of cystitis, accommodated 24 genes that had different sequences when compared to other strains. Therefore, a large group of genes among four strains of ST95 were identified that represented variability among their sequences when compared among four ST95 strains. Hence another alternative approach was put forward to identify specific candidate genes for SNP analysis.

4.2.b. *In silico* analysis of 46 genomes of *E. coli* strains including STC95 strains, based on recombination and mutation events to identify candidate genes for SNP analysis discerning ST95 strains according to specific human/avian host origin

As a result of genome comparison of 46 *E. coli* genomes, many different genes were detected as mentioned above (4.2.a) that showed variation in their sequences among ST95 strains. Thus we used an additional alternative method for the identification of genes useful for SNP analysis. As the genes undergoing frequent recombination could not be used for SNP analysis, further gene differentiation of genes was based on recombination and mutation events within the genes to identify genes for SNP analysis.

In this method, individual gene trees for all the 2,169 genes identified earlier were obtained by using Phylogenetic estimation using Maximum Likelihood (phyML). A single tree of an individual gene was later compared with the additional Parsimony tree obtained using PHYLIP (Phylogeny Inference Package), which was reconstructed from the concatenation of all the genes. The recombination events among all the gene trees were then detected based on *p*-values obtained by AU test (Approximately unbiased) and recombination was detected using MaxChi (Maximum Chi square test) and Pairwise Homoplasy index (PHI). Genes that had undergone recombination were rejected and only those genes that had undergone some mutations were selected. As a result of this approach, 36 genes were selected (Table 8) and considered to be candidate genes for SNP analysis.

Table 8: Candidate genes identified to have mutations and no recombination history. These genes were identified by the analysis of 46 genomes with a method detecting recombination and mutation events within a gene.

No.	Gene	Gene product
1	<i>gapA</i>	Glyceraldehyde-3-phosphate dehydrogenase A
2	<i>rplB</i>	50S ribosomal protein L2
3	<i>torR</i>	DNA-binding transcriptional regulator TorR
4	<i>rsxA</i>	Na(+)-translocating NADH-quinone reductase subunit E
5	<i>sodB</i>	Superoxide dismutase
6	<i>lrp</i>	Leucine-responsive transcriptional regulator
7	<i>pal</i>	Peptidoglycan-associated outer membrane lipoprotein
8	<i>rpsG</i>	30S ribosomal protein S7
9	<i>rplE</i>	50S ribosomal protein L5
10	<i>rplF</i>	50S ribosomal protein L6
11	<i>ygaP</i>	Conserved hypothetical protein

No.	Gene	Gene product
12	<i>ompX</i>	Outer membrane protein X
13	<i>spy</i>	Periplasmic protein
14	<i>sixA</i>	Phosphohistidine phosphatase
15	<i>dksA</i>	RNA polymerase-binding transcription factor
16	<i>yhbC</i>	Hypothetical protein
17	<i>yaiY</i>	Conserved hypothetical protein
18	<i>metJ</i>	Transcriptional repressor protein MetJ
19	<i>rbfA</i>	Ribosome-binding factor A
20	<i>rplP</i>	50S ribosomal protein L16
21	<i>yraN</i>	Hypothetical protein
22	<i>apaG</i>	ApaG protein
23	<i>dinI</i>	DNA damage-inducible protein I
24	<i>rnpA</i>	Ribonuclease P
25	<i>ydhD</i>	Hypothetical protein
26	<i>ybaB</i>	Hypothetical protein
27	<i>smpA</i>	Hypothetical protein
28	<i>ybjQ</i>	Hypothetical protein
29	<i>clpS</i>	ATP-dependent Clp protease adaptor protein ClpS
30	<i>yjfN</i>	Hypothetical protein
31	<i>yciS</i>	Conserved hypothetical protein
32	<i>yciI</i>	YciI-like protein
33	<i>fis</i>	DNA-binding protein Fis
34	<i>rpoZ</i>	DNA-directed RNA polymerase subunit omega
35	<i>minE</i>	Cell division topological specificity factor MinE
36	<i>yaiB</i>	Hypothetical protein

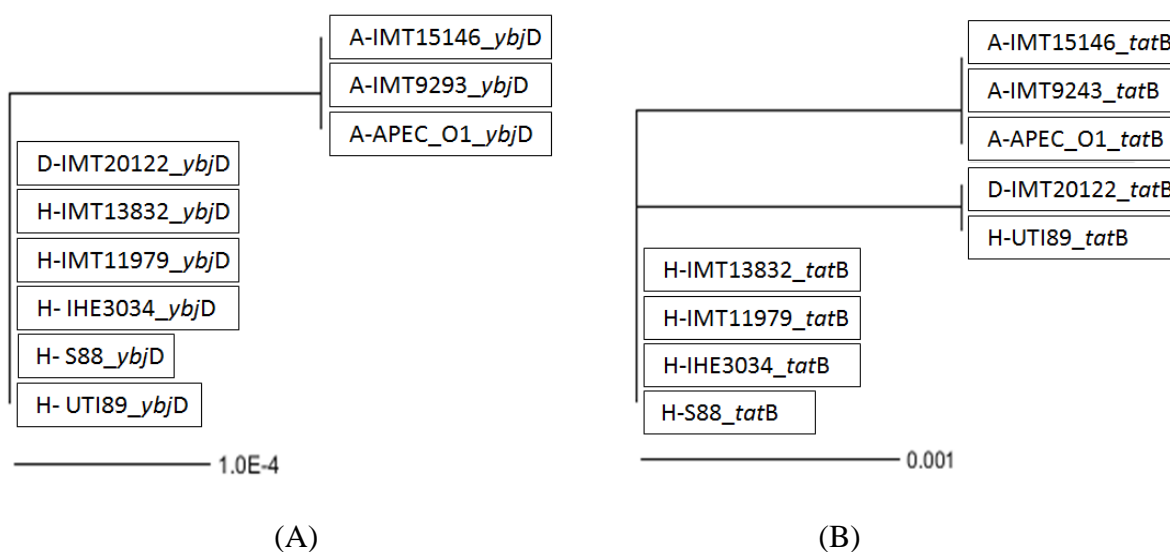
When the sequences of all these 36 genes were compared within the four strains of ST95 (mentioned earlier), the sequence identity for each of the gene among these four strains was observed to be 100%, illustrating no variability in the sequences among ST95 strains for all the genes. Hence, these 36 genes could not be used to differentiate the APEC_O1 strain from the other three human strains of ST95. Therefore, the selection of genes by identifying mutation and recombination events within genes could not identify candidate genes for SNP analysis.

4.2.c. Identification of genetic markers by analyzing the MCG of nine genomes of strains of ST95

As the genome analysis mentioned earlier could not point out candidate genes for SNP analysis, another alternative method was put forward to identify candidate genes for SNP typing. The MCG of nine strains of ST95 was analyzed to contain 1,300 genes. Among these nine strains, the sequences of four ST95 strains (APEC_O1, S88, IHE3034 and UTI89) were available at

NCBI and the other additional five strains (two strains [IMT11979, IMT13832] isolated from human, two strains isolated from birds [IMT9243, IMT15146], and one strain of canine origin [IMT20122]) that belong to ST95 were sequenced by our institute).

As a result of this analysis twelve genes were identified (Table 9) that could segregate the strains isolated from birds on different branch of multiple sequence alignment tree from the strains isolated from human and dog, based on their sequence variation as shown in Figure 5.



(A = Avian, H = Human, D = Dog)

Figure 5: Multiple Alignment trees developed using Ridom Seqsphere representing trees for two genes (A = *ybjD*) and (B = *tatB*) separating strains of human from strains of avian origin. (Two strains isolated from humans [IMT11979, IMT13832], two strains of avian origin [IMT9243, IMT15146] and one dog strain IMT20122) that belong to ST95 were sequenced by our institute.

All 12 genes (Table 9) were amplified from 24 strains of human and 24 strains of avian origin of ST95 available at the time of analyses and were sequenced and analysed for SNPs using MEGA software (Molecular Evolutionary Genetics Analysis Version 5.05) (Tamura et al. 2011).

Table 9: Candidate genes selected for SNP analysis.

No.	Genes	Size (bp)	Gene product
1	<i>ybjD</i>	600	Hypothetical protein with nucleoside triphosphate hydrolase domain
2	<i>tatB</i>	200	Sec- independent translocase
3	<i>mazG</i>	792	Nucleoside triphosphate pyrophosphohydrolase

No.	Genes	Size (bp)	Gene product
4	<i>hemH</i>	963	Acetyl esterase
5	<i>hycD</i>	924	Membrane spanning protein of formate hydrogenase
6	<i>hisJ</i>	723	Histidine/lysine/arginine/ornithine transporter subunit
7	<i>lon</i>	240	DNA binding ATP dependent protease La
8	<i>rbn</i>	873	Ribonuclease BN
9	<i>rmfA</i>	621	Translocating NADH quinone reductase
10	<i>yehN</i>	354	Hypothetical protein
11	<i>yciM</i>	117	Tetraricopeptide repeat protein
12	<i>yniC</i>	669	2- deoxyglucose-6- phosphatase

(The gene size in base pair represents the sequences within the gene that possessed specific SNPs differentiating the three genomes of strains of avian origin from five genomes of strains isolated human as shown in Fig. 5. These fragments of genes were amplified from 24 strains of human and 24 strains of bird using PCR. The strain APEC_O1 (NC_008563.1) was used as a reference for procuring sequences from NCBI for designing primers accordingly).

We found that all the twelve genes that could separate the nine fully sequenced ST95 strains (mentioned earlier) based on specific SNP positions as shown in Figure 5 were unsuccessful to discern the strains isolated from humans and birds with respect to their host when a larger set of strains (total 48 strains, 24 strains of human and 24 strains of avian) was analysed. Since no host specific SNP was found, the strains from humans and avian hosts appear on the same branch of the multiple sequence alignment tree (Figure 6). Out of all twelve genes that were analysed for host specific SNPs none of the genes represented host specific SNPs. Hence, it can be concluded that out of the 1,300 genes that were initially analysed for identifying candidate genes for SNP typing, none represented promising SNPs that could discern the ST95 strains with respect to their host. These data indicate that the MCG of human and animal ST95 strains does not provide any host specific genetic difference. Thus from this analysis we presume these strains to be zoonotic.



Figure 6: Multiple Alignment trees developed using Ridom Seqsphere representing trees for two genes (A) *ybjD* and (B) *tatB* that could not separate larger group of strains of human and avian origin as these strains appear on the same branch. Strains isolated from humans are marked in boxes.

4.2.d. SNP analysis of *E. coli* fimbrial adhesin gene

During the bioinformatics approach to identify suitable genes for SNP analysis from the MCG we complemented the ongoing analysis with SNP typing of a gene that has already been suggested suitable for discerning ExPEC groups. This gene, the fimbrial adhesin gene *fimH*, was therefore selected for further analyses.

FimH is part of a chromosomal operon and its gene product (adhesin) is responsible for the mannose-specific or receptor-specific binding of this type of fimbriae. In particular, specific *fimH* alleles expressed by many uropathogenic *E. coli* exhibit higher affinities for mono-mannose receptors than most intestinal commensal isolates (Pouttu et al. 1999). This higher mono-mannose affinity increases bacterial tropism for uroepithelium and thus promotes bladder colonization (Sokurenko et al. 1998). This variation in FimH adhesin enhances its binding capacity to targets such as laminin, collagen and fibronectin as well as to different mannose derivatives (Bouckaert et al. 2005). Pathoadaptive mutations result in allelic variations of the FimH adhesin subunit ((Weissman et al. 2006, Weissman et al. 2007, Weissman et al. 2012). These mutations are important in tissue tropism determining fine sugar specificity of these fimbriae (Sokurenko et al. 1999, Oelschlaeger et al. 2002). By investigating such variability by SNP typing we hypothesized that it would be possible to identify distinctive pathoadaptive mutations within a given phylogenetic group such as ST95 with regard to the specific host.

Our SNP typing revealed that the 108 strains (50 strains isolated from humans and 58 strains of avian origin) of ST95 represented 14 different *fimH* alleles (Table 10). However, as this table outlines even strains isolated from human and birds represented identical *fimH* alleles, thus indicating that a particular allelic variation is not related to a particular host. Identical *fimH* alleles were observed in the strains of avian and human origin. Therefore, *fimH* SNP typing also does not prove to be a suitable genetic marker to differentiate strains of ST95 with respect to their host.

Table 10: FimH allele types (based on nucleotide sequences) and frequency among ST95 ExPEC strains

ST95 ExPEC strains positive for the allele				
FimH allele	Strains isolated from humans	Strains isolated from birds	Total Number	Percentage (%)
I	14	29	43	39.09
II	14	9	23	20.91
III	3	7	10	9.09
IV	7	3	10	9.09
V	5	0	5	4.55
VI	1	2	3	2.73
VII	2	1	3	2.73
VIII	0	3	3	2.73
IX	1	1	2	1.82

X	2	0	2	1.82
XI	0	1	1	0.91
XII	0	1	1	0.91
XIII	1	0	1	0.91
XIV	0	1	1	0.91
Total	50	58	108	100.00

Some alleles were observed more frequent within the ST95 strains (Table 10) creating an interest to investigate the phenotypic characters of strains of these predominant allelic variants.

4.3. Cloning of prominent *fimH* alleles

As mentioned above 14 different *fimH* alleles were obtained in the ST95 strains of different pathotypes. Therefore we decided to investigate the phenotypic characters of some strains from predominant allelic variants with respect to the *fimH* gene. With this regard, cloning of the *fim* operon (~8.7 kb) from some of the strains (Table 5 see 3.2.c.iii) belonging to the predominant allelic groups (I, II, III, IV, V) using TOPO cloning and also direct cloning in pET-45b(+) vector was performed. However, successful clones were not obtained even when the cloning procedure was performed several times. The successful cloning and transformation of *fim* negative strains with *fim* clones could have helped us to understand the importance of the *fim* operon and particularly the role of pathoadaptive mutations within FimH. FimH plays a role in the adhesion of bacteria and in the early stages of biofilm formation by bacteria and genetic variability and specific polymorphisms within FimH facilitates bacterial adaptation towards host environment and thereby leading to infection (Sokurenko et al. 1998).

4.4. Identification of phenotypic characteristics in ST95 strains that could segregate this group of strains according to their human/avian host respectively

As the second aim of this study was to identify specific phenotypes among the strains of ST95, different phenotypic characters like adhesion and invasion capabilities, biofilm formation capabilities, long-term biofilm formation, haemagglutination activity and serum sensitivity in chicken and human serum were investigated to identify certain host specific phenotypes.

4.4.a. Adhesion and invasion assay

The adhesion and the invasion capabilities of thirteen strains isolated from humans (six NMEC, six UPEC strains and one commensal strain) and eight strains of avian origin (six APEC strains and two commensal strains) out of 116 ST95 strains were initially investigated in HEK293-T (Human embryonic kidney) cell line and MDCK-1 (Madin-Darby Canine Kidney-1) cell line.

The strains selected for adhesion and invasion assay had similar clonal nature as observed by their ST and similar virulence gene pattern and macro-restriction pattern of the whole genome (Pulsed-field gel electrophoresis) as determined before the start of this thesis (Ewers, Wieler et al., unpublished data). The selection of the cell line considered the background knowledge that UPEC strains adhere to bladder epithelial cells and ascend towards kidney; hence we investigated if ST95 strains of human origin representing different pathotypes adhere to the kidney cells and if the same phenotype is also demonstrated by ST95 strains of avian origin. The dog cell lines was also used so as to investigate if the ST95 strains can adhere to and invade kidney cells from a host other than humans and birds as recently two ST95 strains had been isolated from dogs (Ewers, Wieler et al., unpublished data).

Adhesion in HEK293-T cell line and MDCK-1 cell line

As mentioned, strains isolated from humans and birds were assessed for their ability to adhere to HEK293-T cell line. The adhesion profile of strains isolated from human and birds is shown in Figure 7.

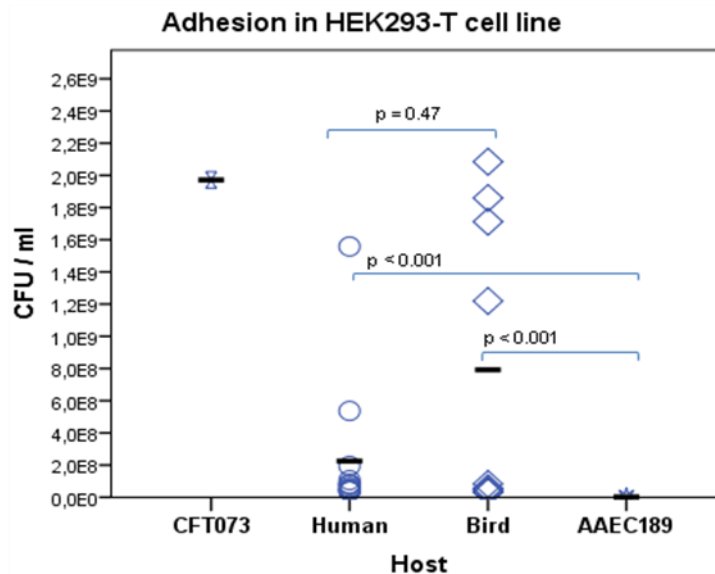


Figure 7: Adhesion capabilities of 13 strains isolated from humans (6 NMEC, 6 UPEC strains and 1 commensal strain) and 8 strains isolated from birds (6 APEC and 2 commensal strains) on HEK cells after 3 hrs of incubation with HEK cell line in mannose-free media. Black bar represents the mean. CFT073 (human UPEC strain) was used as a positive control for adhesion and AAEC189 (*fim*-negative K-12 derivative) as negative control. The assay was performed three times in triplicates.

As outlined in Fig. 7, strains isolated from humans and birds represented adherence capabilities towards human embryonic kidney cell line as compared to the negative control. Strains isolated from humans, particularly UPEC strains are known to adhere to the kidney cells after ascending from the bladder. Interestingly, some of the strains isolated from birds represented more adherence than the strains of human origin. However, from our results it can be concluded that the strains isolated from birds were capable of adhering to the human kidney cell line as that of the strains isolated from humans and as there was no significant difference in their adherence properties (Mann Whitney U test, $U = 47$, $p = 0.47$). Notably, two ST95 commensal strains isolated from birds that had similar virulence genes profile and PFGE pattern as that of the APEC strains of ST95 demonstrated high adhesive properties towards human cell line apart from the pathogenic avian strains.

Strains isolated from humans and birds (used in earlier experiments with the HEK cell line) were also assessed for their ability to adhere to other host cells like Madin-Darby Canine Kidney (MDCK-1) cell line from dog. The adhesion profile of strains of humans and birds towards MDCK-1 cells is shown in Figure 8.

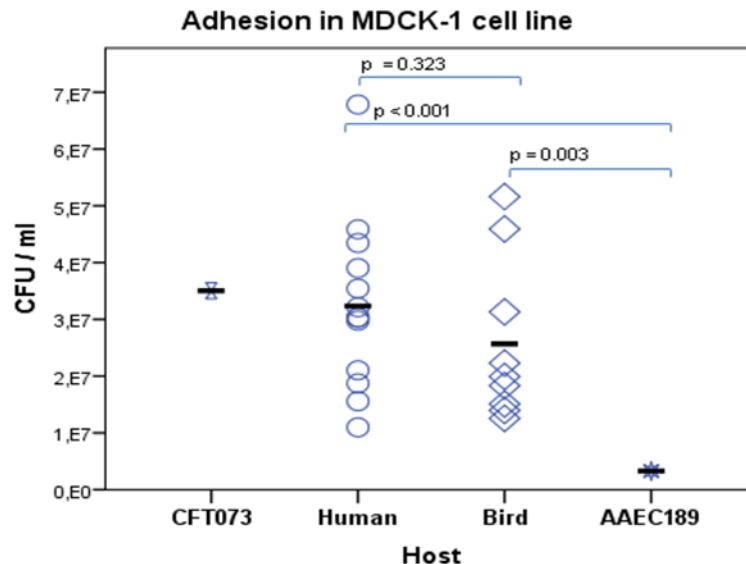


Figure 8: Adhesion capabilities of 13 strains isolated from humans (6 NMEC, 6 UPEC strains and 1 commensal strain) and 8 strains of avian origin (6 APEC and 2 commensal strains) on MDCK-1 cells after 3 hrs of incubation with MDCK-1 cell line in mannose-free media. Black bar represents the mean. CFT073 (human UPEC strain) was used as a positive control for adhesion and AAEC189 (*fim*-negative K-12 derivative) as negative control. The assay was performed three times in triplicates.

Both the group of strains isolated from humans and birds demonstrated adherence towards MDCK-1 cells. However, it was further evident that the strains of avian origin were capable of adhering to MDCK-1 cell line as strong as the strains from humans that are known to adhere to kidney cells. There was no significant difference in the adherence properties of avian and human strains ($U = 42$, $p = 0.323$). The adhesion capabilities of both the group of strains towards cell line from another host indicate that there could be zoonotic properties among ST95 strains and these strains could be successful pathogens to cause infections in a host other than humans and birds.

Invasion capabilities in HEK-293 cell line and MDCK-1 cell line

The invasion capability of the same strains that were used in the adhesion assay was also investigated in human cell line in accordance with the gentamicin protection assay, where the extracellular bacteria are killed by the bactericidal effect of antibiotic and the invasive intracellular bacteria escape the killing effect of antibiotic (Elsinghorst 1994). All strains used in the experiment were unable to grow on agar plates after 1.5 hrs of incubation in DMEM containing gentamicin at a working concentration of 50 mg/ml. Therefore they could be investigated in the gentamicin protection assay. The invasive abilities of both the groups of strains are given in Figure 9.

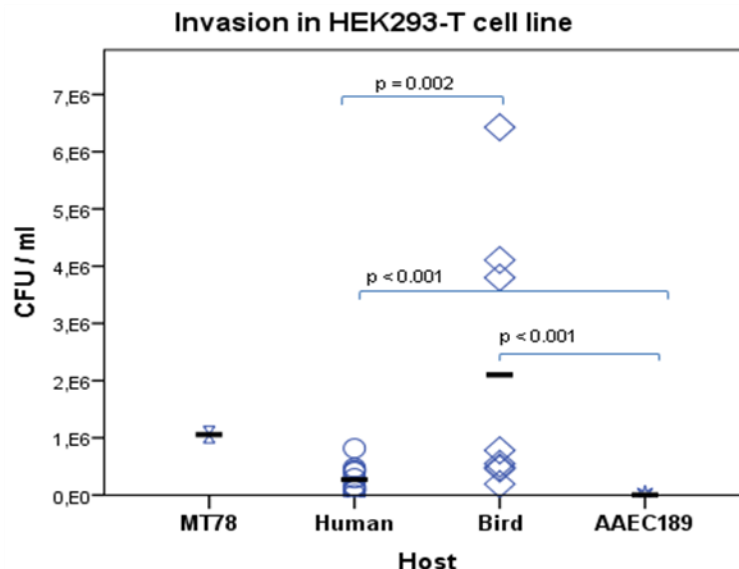


Figure 9: Invasion capabilities of 13 strains isolated from humans (6 NMEC, 6 UPEC strains and 1 commensal strain) and 8 strains of avian origin (6 APEC strains and 2 commensal strains) in HEK cells after 3 hrs of incubation with HEK cell line in mannose-free media and additional incubation for 1.5 hrs in media containing gentamicin (50 mg/ml) to kill the surface adherent bacteria after removing the earlier media used during

the incubation period of 3 hrs. Black bar represents the mean. MT78 was used as a positive control for invasion assay and AAEC189 (*fim*-negative K-12 derivative) as negative control. The assay was performed three times in triplicates.

Both the group of strains isolated from humans and birds represented invasion capabilities in HEK cells. However, the invasion experiments indicated that the strains of avian origin were potentially more invasive towards human cell line as compared to the human strains ($U = 11$, $p = 0.002$). Some of the strains of avian origin (one APEC strain and two ST95 commensal strains that also illustrated higher adherence capabilities and possessed similar virulence gene pattern and PFGE profile as that of the APEC strains) demonstrated high invasive phenotype towards human cells.

The invasion capabilities of the same strains from humans and birds (those used with HEK cell line) were also studied in MDCK-1 cell line using gentamicin protection assay and the results are shown in Figure 10. As compared to the invasion capabilities of these strains towards HEK cell line both the group of strains from humans and avian hosts were less invasive in MDCK-1 cells.

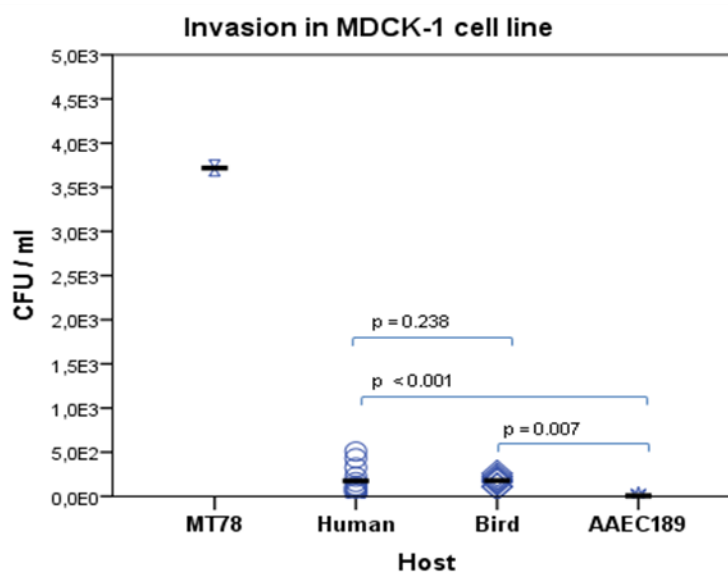


Figure 10: Invasion capabilities of 13 strains isolated from humans (6 NMEC, 6 UPEC strains and 1 commensal strain) and 8 strains of avian origin (6 APEC and 2 commensal strains) in MDCK-1 cells after 3 hrs of incubation with MDCK-1 cell line in mannose-free media and additional incubation for 1.5 hrs in media containing gentamicin (50 mg/ml) to kill the surface adherent bacteria after removing the earlier media used during the incubation period of 3 hrs. Black bar represents the mean. MT78 was used as a positive control for invasion assay and AAEC189 (*fim*-negative K-12 derivative) as negative control. The assay was performed three times in triplicates.

As compared to the positive control (MT78), the ST95 strains isolated from humans and birds were less invasive as illustrated in the Figure 10. The strains of human and avian origin did not differ much in their invasion capabilities towards MDCK-1 cells indicating no significant difference ($U = 35, p = 0.238$) in their invasion properties towards MDCK-1 cell line.

4.4.b. Biofilm formation assay

4.4.b.i. 96-well microtiter plate method

As the adhesion and invasion experiments did not show much of a difference among ST95 strains from humans and birds, we analyzed the ability of these strains to form biofilms. For this static 96-well microtiter plate biofilm assays were utilized, using two different media i.e. nutrient medium (LB) and M63 minimal medium. In total, 116 ST95 strains available at our institute were investigated. We observed media-dependent biofilm forming capability of strains from humans and avian hosts. In LB media, only few strains (avian - 16% and human - 9%) produced biofilms after 24 hrs of incubation. After 48 hrs of incubation in LB media 23% of strains from avian host and 18% of strains from humans were able to form biofilms. However, in case of M63 media a higher number of strains (30% of avian origin and 35% of human strains) demonstrated biofilm formation after 24 hrs and after 48 hrs in M63 media the number of strains producing biofilm was highest (64% of strains from birds and 69% of strains from humans – 69%, see Figure 11). The strains that formed biofilms belonged to different pathotypes and there were no specific pathotype of strains representing high biofilm forming capacity in both the groups.

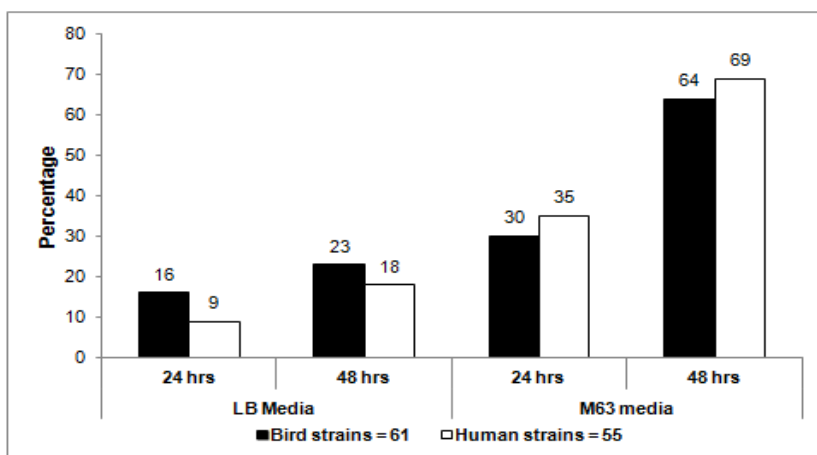


Figure 11: Biofilm formation capabilities (in percentage) of 61 *E. coli* strains isolated from birds and 55 *E. coli* strains isolated from humans in two different media (LB and M63) conditions after 24 hrs and 48 hrs of incubation. For each strain the experiment was performed three times in triplicate.

The results from the microtiter biofilm assay also illustrated that the ST95 strains from avian host demonstrated similar biofilm forming capacity under different media conditions as those strains isolated from humans (Figure 11).

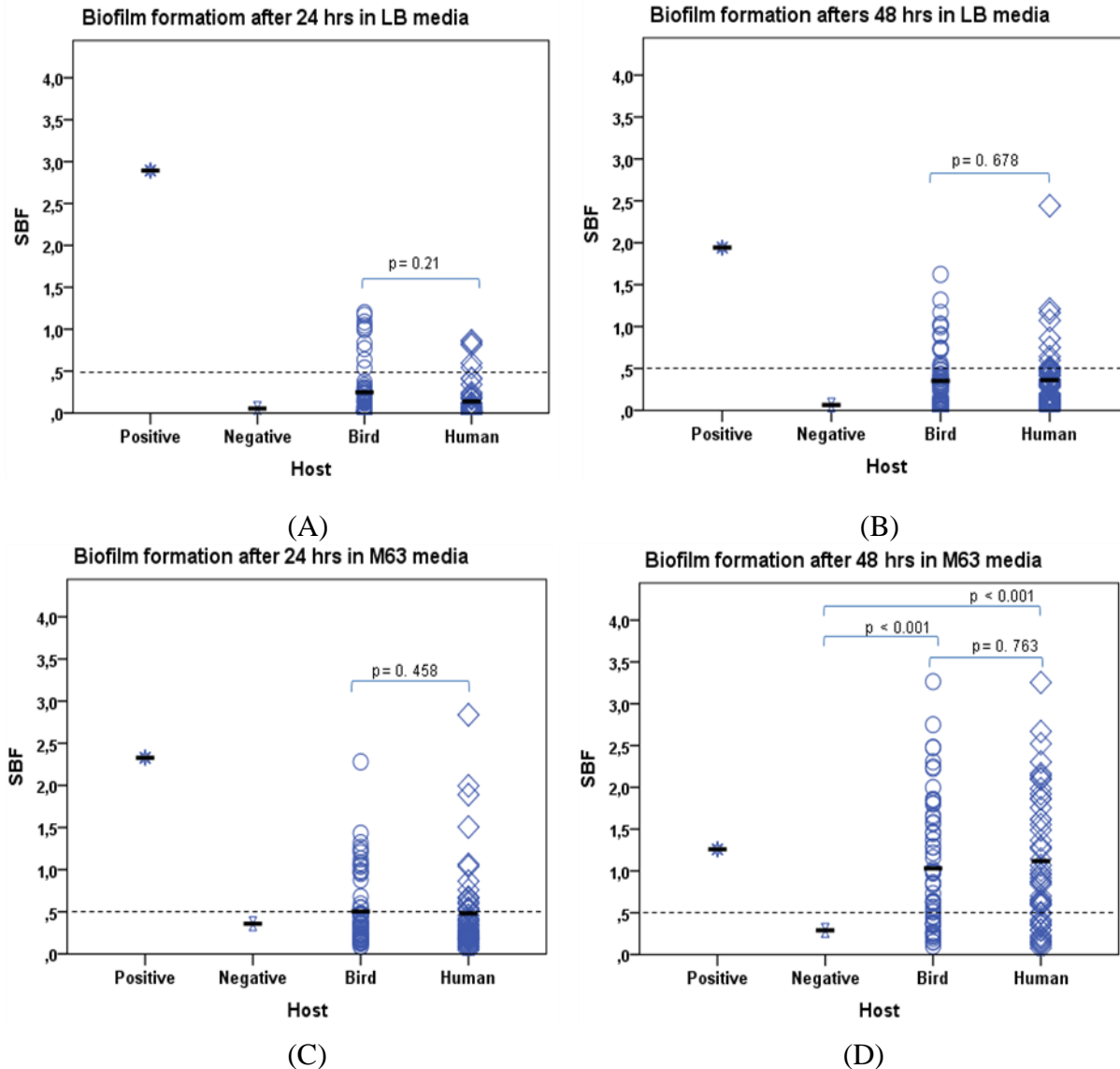


Figure 12: Scatter plot showing SBF (specific biofilm formation) capacities of 55 strains from humans and 61 strains from birds. Figure (A) and (B) show SBF in LB media after 24 and 48 hrs of incubation and figure (C) and (D) represent SBF in M63 media after 24 and 48 hrs of incubation. *E. coli* strain W3110 (*E. coli* K-12 derivative) was used as a positive control and AAEC189 (*E. coli* K-12 derivative) served as a negative control. Strains above reference line (0.5) were considered as biofilm forming strains. The black bar indicates the mean values. The SBF for each strain was determined by three different experiments in triplicate.

Biofilm formation by both groups of strains in M63 media was observed to be higher after 48 hrs than after 24 hrs of incubation (Figure 12). Thus, considering all the results, significant

difference was proven between the non-biofilm forming AAEC189 strain (negative control) and the strains isolated from humans and birds (strains from birds and from humans, both $p < 0.0001$) in Figure 12 (D). However, most importantly there was no difference in the biofilm forming capacity between strains isolated from humans and birds after 48 hrs in M63 media ($p = 0.763$). Likewise, no differences were observed between strains isolated from birds and strains from humans under different media conditions and incubation time indicating that the strains isolated from birds represented a similar phenotypic property of biofilm formation as that of strains isolated from humans under different sets of conditions.

4.4.b.ii. Long term biofilm formation

The expression of curli and cellulose via long term biofilm formation on Congo red plates was also investigated among the ST95 strains that were earlier investigated for biofilm forming capacities.

Curli fimbriae and polysaccharide cellulose are known to promote bacterial attachment to the host intestinal epithelium and biofilm development. With this background we investigated whether the ST95 strains isolated from humans and avian hosts demonstrate differences in the expression of curli fimbriae and cellulose under different temperatures (28°C and 37°C) and different media conditions (with and without salt). Different colony morphotypes used as a reference for the determination of curli and cellulose expression are shown in Figure 13.

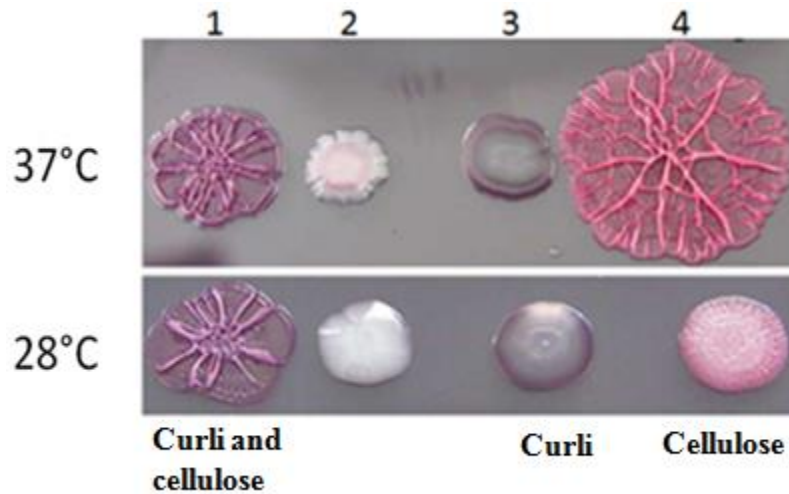
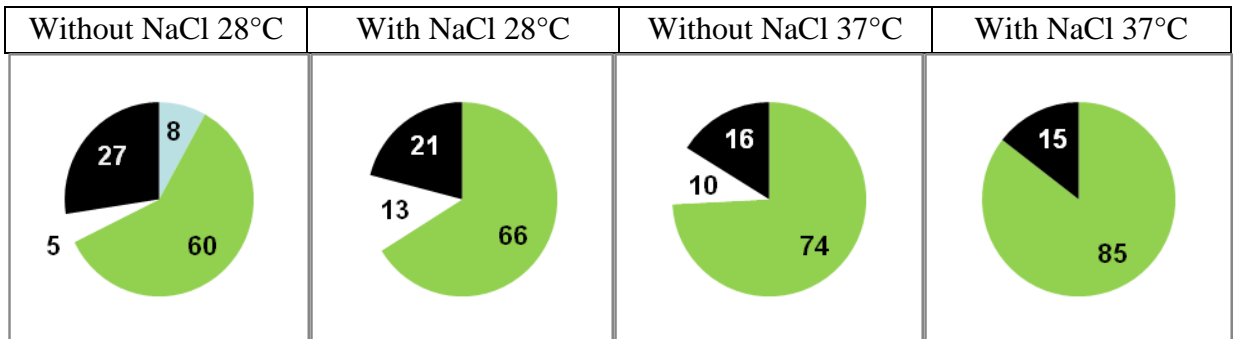
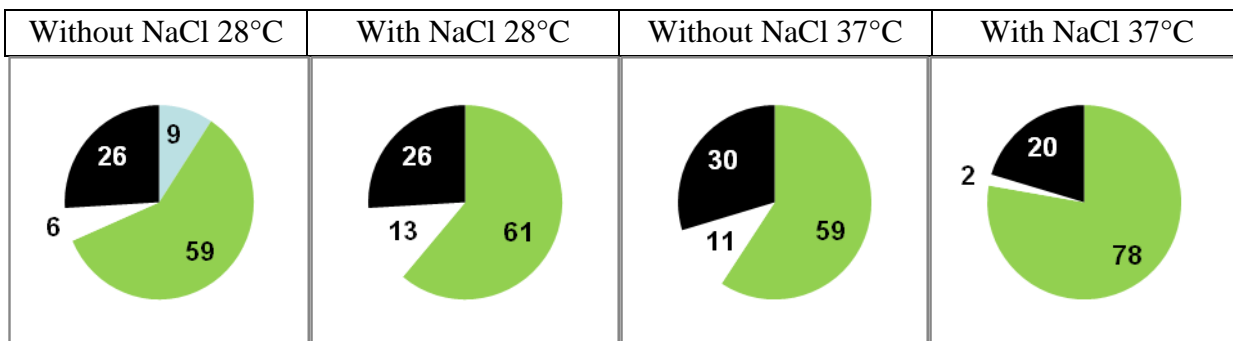


Figure 13: Different colony morphotypes used as a reference representing expression of curli and cellulose (Bokranz et al. 2005). 1- Expression of curli and cellulose, 2- No expression of curli and cellulose, 3- Expression of curli, 4- Expression of cellulose (IMT26949-UPEC strain from dog was used as a control for expression of curli and cellulose).

To our observation, we found that there was differential expression of curli and cellulose by the strains isolated from humans and birds as shown in Figure 14.



61 strains isolated from birds



55 strains isolated from humans

1 Expression of curli and cellulose 2 No expression of curli and cellulose 3 Expression of curli 4 Expression of cellulose

Figure 14: Expression of curli and cellulose in percentage by 61 strains from birds and 55 strains from humans under different media (with and without salt) and temperature (28°C and 37°C) conditions on Congo red containing agar plates after 4 days of incubation.

Both the group of strains isolated from humans and avian hosts demonstrated almost similar patterns of curli and cellulose expression under different temperature and media conditions. They showed the expression of curli and cellulose only at 28°C on plates without salt (human - 9%, avian - 8%) while this morphotype was not observed at 37°C on plates with salt and without salt. The expression of only curli morphotype by both the groups was observed at 28°C on plates with salt and without salt but, interestingly at 37°C the expression of curli was observed only on the plates without salt and not on the plates with salt. Another interesting observation was the expression of cellulose by both the groups of strains irrespective of temperature and media condition. However, most of the strains from both the groups neither showed expression of curli and cellulose together nor separately. Thus, the results indicate that the presence and absence of salt influences the expression of curli and cellulose at the same temperature conditions and this differential expression behavior was observed to be same in the ST95 strains isolated from humans and birds. In other words, the strains from birds show similar expression pattern of curli and cellulose as that of the strains isolated from humans.

The strains of avian origin expressed curli and cellulose morphotypes only at lower temperature of 28°C, which is a known characteristic of strains, isolated from humans particularly UTI isolates, giving the impression that at lower temperature the strains of avian origin may have a similar expression behavior for curli and cellulose as that of human-derived strains.

When the culture conditions were assessed against the different morphotypes of the colonies representing expression of curli and cellulose using multinomial logistic regression method, it could be concluded that, the probability of expression of curli morphotypes by strains isolated from humans and birds increases ($p = 0.021$) (Table 11) when the culture condition was set at 28°C and the plates used were without NaCl.

Table 11: Cross tabulation between the culture conditions and expression patterns.

Parameter Estimates									
Morphology of the Colony ^a		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp (B)	
								Lower Bound	Upper Bound
Curli and cellulose	Intercept	-22,698	,338	4499,839	1	,000			
	[Condition=1,00]	20,766	,000	.	1	.	1044153879	1044153879	1044153879
	[Condition=2,00]	,267	8634,860	,000	1	1,000	1,306	,000	^b
	[Condition=3,00]	,238	8533,524	,000	1	1,000	1,268	,000	^b
	[Condition=4,00]	0 ^c	.	.	0
Only curli	Intercept	-1,558	,246	40,112	1	,000			
	[Condition=1,00]	,758	,328	5,356	1	,021	2,134	1,123	4,055
	[Condition=2,00]	,550	,333	2,723	1	,099	1,733	,902	3,331
	[Condition=3,00]	,460	,334	1,889	1	,169	1,583	,822	3,049
	[Condition=4,00]	0 ^c	.	.	0
Only cellulose	Intercept	-4,554	1,005	20,522	1	,000			
	[Condition=1,00]	2,112	1,092	3,741	1	,053	8,261	,972	70,183
	[Condition=2,00]	2,958	1,044	8,021	1	,005	19,257	2,487	149,125
	[Condition=3,00]	2,682	1,052	6,500	1	,011	14,615	1,859	114,885
	[Condition=4,00]	0 ^c	.	.	0

a. The reference category is: No curli and cellulose.

[The table represents the results of multinomial logistic regression analysis using Statistical Package for the Social Sciences (SPSS, version 10.0) when the culture conditions (temperature of 28°C and 37°C and plates with and without salts) were checked against the colony morphotypes representing expression of curli and cellulose either together or separately].

However, when the strains isolated from humans and birds were assessed by including the variable host into a multivariable model the results did not change as shown in Table 12.

Table 12: Cross tabulation of culture conditions, expression patterns and the host.

Parameter Estimates									
Morphology of the Colony ^a		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp (B)	
								Lower Bound	Upper Bound
Curli and cellulose	Intercept	-22,829	,472	2340,604	1	,000			
	[Condition=1,00]	20,770	,000	.	1	.	1047704287	1047704287	1047704287
	[Condition=2,00]	,268	8638,946	,000	1	1,000	1,308	,000	^b
	[Condition=3,00]	,239	8535,905	,000	1	1,000	1,270	,000	^b
	[Condition=4,00]	0 ^c	.	.	0
	[Host=1,00]	,273	,666	,168	1	,682	1,314	,356	4,848
	[Host=2,00]	0 ^c	.	.	0
Only curli	Intercept	-1,738	,274	40,213	1	,000			
	[Condition=1,00]	,764	,329	5,413	1	,020	2,148	1,128	4,089
	[Condition=2,00]	,555	,334	2,755	1	,097	1,741	,905	3,353
	[Condition=3,00]	,463	,335	1,911	1	,167	1,590	,824	3,067
	[Condition=4,00]	0 ^c	.	.	0
	[Host=1,00]	,364	,228	2,544	1	,111	1,439	,920	2,252
	[Host=2,00]	0 ^c	.	.	0
Only cellulose	Intercept	-4,676	1,023	20,906	1	,000			
	[Condition=1,00]	2,116	1,092	3,756	1	,053	8,297	,976	70,515
	[Condition=2,00]	2,961	1,045	8,037	1	,005	19,322	2,494	149,675
	[Condition=3,00]	2,685	1,052	6,511	1	,011	14,655	1,864	115,235
	[Condition=4,00]	0 ^c	.	.	0
	[Host=1,00]	,254	,368	,479	1	,489	1,290	,627	2,651
	[Host=2,00]	0 ^c	.	.	0

a. The reference category is: No curli and cellulose.

[The table represents the results of multinomial logistic regression analysis using Statistical Package for the Social Sciences (SPSS, version 10.0) when the culture conditions (temperature of 28°C and 37°C and plates with and without salts) were checked against the colony morphotypes representing expression of curli and cellulose either together or separately expressed by avian (61 strains) and human (55 strains) of ST95].

Therefore, the expression of curli morphology at 28°C and without NaCl was found to be independent of the host. Alternatively, there was no difference in the expression of curli at 28°C on plates without NaCl among strains isolated from humans and birds. Hence, indicating that the ST95 strains of avian hosts have similar curli expression behaviour as that of the strains isolated from humans making them indistinguishable from human strains in this respect.

4.4.c. Haemagglutination assay

Haemagglutination assay was performed with chicken erythrocytes in the presence and absence of D-mannose. Agglutination of erythrocytes by bacteria is linked to the expression of different adhesins. The expression of type 1 fimbriae was assessed by performing haemagglutination assays in the presence and absence of mannose for ST95 strains that were used in earlier experiments. . The results of haemagglutination assays are shown in Figure 15.

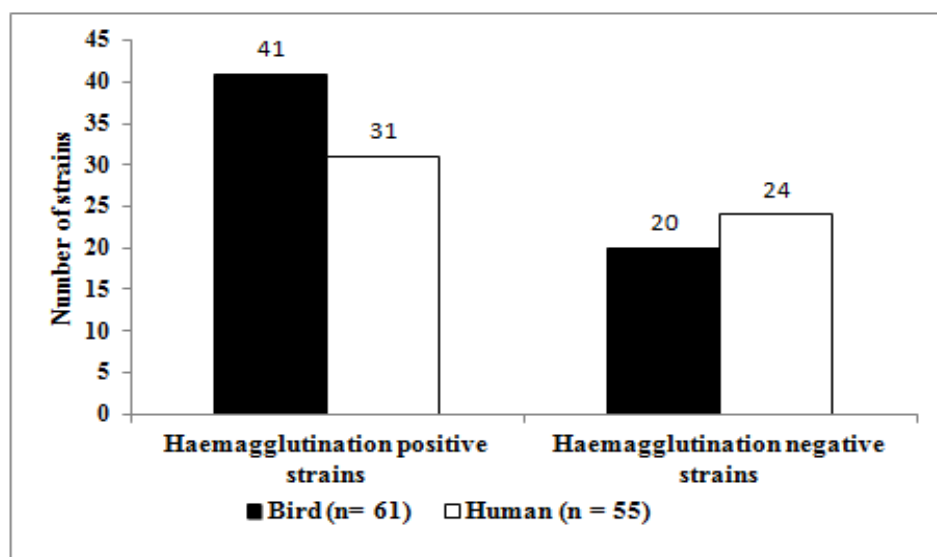


Figure 15: Haemagglutination results with chicken erythrocytes for 61 *E. coli* strains of birds and 55 *E. coli* strains of humans in the absence of mannose. The experiment has been repeated three times in triplicate. AAEC189 (*fim*-negative K-12 strain) was used as a negative control. When the positive strains (72 strains from humans and birds) were checked for haemagglutination in the presence of 5% D-mannose and chicken erythrocytes none of the strain demonstrated mannose-resistant haemagglutination (MSHA). No significant difference was observed between the strains isolated from humans and strains of avian origin that were positive for haemagglutination $\chi(1) = 1.446$, $p = 0.229$.

Table 13: Host * Haemagglutination Cross tabulation

		Haemagglutination		Total	
		Haemagglutination positive	Haemagglutination negative		
Host	Human	Count	31	24	55
		% within Host	56,4%	43,6%	100,0%
		% within Haemagglutination	43,1%	54,5%	47,4%
		% of Total	26,7%	20,7%	47,4%
	Bird	Count	41	20	61
		% within Host	67,2%	32,8%	100,0%
		% within Haemagglutination	56,9%	45,5%	52,6%
		% of Total	35,3%	17,2%	52,6%
Total	Count	72	44	116	
	% within Host	62,1%	37,9%	100,0%	
	% within Haemagglutination	100,0%	100,0%	100,0%	
	% of Total	62,1%	37,9%	100,0%	

[Cross tabulation prepared by performing chi-square test using Statistical Package for the Social Sciences (SPSS, version 10.0) for 55 strains isolated from humans and 61 strains from birds representing haemagglutination activity against chicken erythrocytes in absence of mannose].

Among the 55 strains isolated from humans, 56.4% were able to agglutinate chicken erythrocytes; among the 61 strains of birds 67.2% were positive for haemagglutination. In total, 62.1% of all the strains tested were positive for haemagglutination (Table 13). When the positive strains were tested for haemagglutination assay with chicken erythrocytes in the presence of D-mannose, no agglutination was observed by any of the strains. This illustrates that ST95 strains do not represent mannose-resistant haemagglutination (MRHA). Our results did not reveal significant differences between the strains isolated from humans and the strains isolated from birds that were positive for haemagglutination $\chi(1) = 1.446$, $p = 0.229$. Thus, it is illustrated that most of the strains from avian hosts also express type 1 fimbriae and show similar agglutination capability like those of the strains isolated from humans and do not serve to be a distinctive phenotype that separate ST95 strains with respect to their host.

4.4.d. Serum assay (chicken and human serum)

As the ST95 strains isolated from humans and birds represented similar phenotypes in the phenotypic assays studied earlier, we investigated their serum resistant phenotype to investigate if they are capable of surviving in the serum from different hosts. The strains that survived in the serum after four hours of incubation were termed as serum resistant strains and those that

did not survive in the serum were termed as serum negative strains. The results from chicken and human serum are shown in Figure 16.

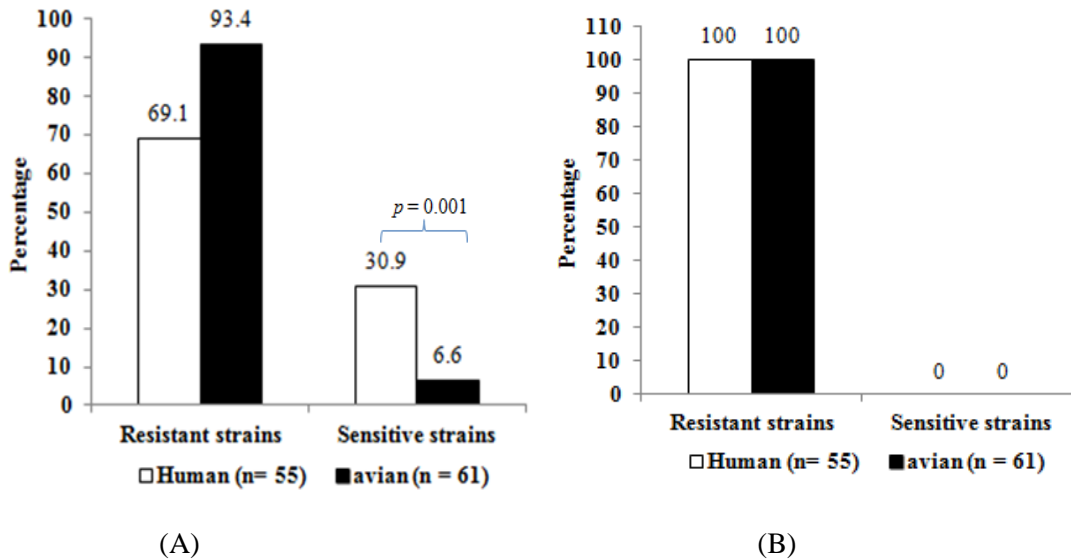


Figure 16: Graph representing percentage resistivity and sensitivity of 55 strains isolated from humans and 61 strains of avian origin in (A) chicken serum and (B) human serum. The strains were incubated in chicken and human serum for four hrs and later diluted in 1 x PBS. 50 μ l of the dilution mixture was dropped in form of small drops on LB agar plates. The colonies were enumerated following day. Each strain was tested two times in triplicates.

To our observation, we found that, out of 55 strains isolated from humans, 17 strains (30.9%) were sensitive to chicken serum and could not survive in chicken serum. Although, seven out of 17 strains isolated from humans that illustrated sensitivity towards chicken serum were commensal strains, interestingly we found that four NMEC, five UPEC and one SePEC strains isolated from humans were sensitive to chicken serum. Another interesting observation was that, out of 38 strains of humans (69.1%) that were resistant to chicken serum, nine strains were commensal and non-pathogenic strains. Out of 61 strains isolated from birds that were investigated for their ability to survive within the chicken serum, 57 strains (93.4%) were resistant to chicken serum and only four strains (6.6%) were sensitive to the lethal effect of chicken serum (Table 14). Interestingly, out of 57 resistant strains, three were non-pathogenic commensal strains and all the four strains that were sensitive to chicken serum were pathogenic APEC strains.

Table 14: Host * Resistance Cross tabulation of chicken serum

			Resistance		Total
			Serum resistant	Serum sensitive	
Host	Human	Count	38	17	55
		% within Host	69,1%	30,9%	100,0%
		% within Resistance	40,0%	81,0%	47,4%
		% of Total	32,8%	14,7%	47,4%
	Bird	Count	57	4	61
		% within Host	93,4%	6,6%	100,0%
		% within Resistance	60,0%	19,0%	52,6%
		% of Total	49,1%	3,4%	52,6%
Total		Count	95	21	116
		% within Host	81,9%	18,1%	100,0%
		% within Resistance	100,0%	100,0%	100,0%
		% of Total	81,9%	18,1%	100,0%

[Cross tabulation prepared by performing chi-square test using Statistical Package for the Social Sciences (SPSS, version 10.0) for 55 strains isolated from humans and 61 strains from birds, representing serum resistivity and serum sensitivity towards chicken serum after four hrs of incubation].

Most of the strains isolated from humans and birds were found to be resistant to chicken serum. However, there was significant difference between the strains of isolated from humans and birds that were sensitive to chicken serum ($\chi(1) = 10.867, p = 0.001$). Thus, it was illustrated that ST95 strains isolated from humans were more sensitive to chicken serum than the strains isolated from birds.

In case of human serum, it was interesting to observe that all the strains from both the groups were resistant to human serum and illustrated considerable growth in human serum. The strains isolated from both the groups (human and avian host) that were sensitive to chicken serum, were also found to be resistant towards human serum.

5. Discussion

Extraintestinal pathogenic *E.coli* (ExPEC) illustrate a wide range of genomic diversity and harbor a broad range of virulence associated genes, often located on pathogenicity islands and other mobile DNA elements (Johnson and Russo 2005, Rodriguez-Siek et al. 2005). The DNA based sequencing method Multilocus sequence typing (MLST) serves as an important tool to unambiguously type strains of a bacterial species in a larger context and can also be used to study the population genetics of microbes. MLST is based on sequence analysis of seven housekeeping genes thereby establishing an allelic profile. The allelic profile is then assigned as a sequence type (ST) (Lau et al. 2008). ExPEC isolates are grouped according to the sequence types based on the allelic profile they exhibit. As observed (Coque et al. 2008, Fam et al. 2011, Guenther et al. 2012, Wu et al. 2012, Habilitation-Ewers 2012) STs 95, 131, 10, 73, 70, 428, 69, 568 and 127 are frequently identified among ExPEC. A previous screening of 1,030 ExPEC strains from various hosts and clinical sources via MLST by our group revealed that these pathogens present a highly diverse group of strains, while some STs appear to be over represented in the population (Habilitation-Ewers 2012). Of these STs, one of the prominent phylotype was ST95, which accumulates highly virulent strains having a wide spectrum of virulence-associated genes as determined by PCR (Ewers, Wieler et al., unpublished data). ST95 comprised strains mainly of human and avian origin. Therefore this ST was designated as “Human-Avian” ST by our group based on previous results, where a large set of strains were determined as non-ST95 strains that were isolated from cattle, swine, dogs and cats, thereby representing limited host range of ST95 strains (Ewers et al unpublished data). As according to the publicly available database (<http://mlst.ucc.ie/mlst/mlst/dbs/Ecoli/> as observed in April 2013) ST95 mainly consists of strains isolated from humans and birds with one strain of canine origin. As most of the strains are from human and avian origin, it supports for the existence of some kind of host limitation of these strains. As *E. coli* is such a variable and heterogeneous bacterial species, both in terms of genomic contents as well as in terms of being commensal or pathogenic (Dobrindt et al. 2003, Ahmed et al. 2008, Tenailon et al. 2010), this limitation of ST95 strains to humans and avian host only make it a remarkable ST interesting for further research.

5.1. Screening of available strains from different animal hosts for sequence type 95

The ST95 consists of strains nearly only from human and avian origin. Thus, to support the host limited nature of ST95, *E. coli* strains from different hosts other than humans and avian sources were screened by PCR for a gene termed “specific for virulent subgroup (*svg*)” that has previously been suggested to specifically identify strains belonging to ST95 (Bidet et al. 2007). Basically, this method was chosen, as it promised to be a rapid and cheap way to screen high numbers of strains belonging to ST95 that would have been later verified by MLST in each case of a positive *svg* result. The screening results of *E. coli* strains from different host with different clinical implications were found to be negative for *svg* gene and could support the hypothesis that the ST95 strains are mostly restricted to human and avian host and only rarely are present in other host species. However, we were able to show that it is not 100% specific as the strain that was positive for this gene, on MLST analysis was assigned ST568 and not as expected to ST95. However, as ST568 belongs to the ST complex 95, at least this strain is highly related to ST95 strains.

A crucial step in understanding bacterial pathogenesis is to discriminate between pathogenic strains from different hosts. From an epidemiological point of view, molecular characterization and differentiation of microorganisms with respect to their pathogenesis and host contribute to the control of infectious diseases. In addition, molecular typing of the bacterial pathogens facilitates differentiation of pathogens in terms of their host specificity, virulence and transmissibility (Feil 2004, Comas et al. 2009). This is of utmost importance in terms of assessing the risk that a particular strain holds for human health. As poultry products serve as veritable food source for humans, contamination of poultry products or infections of humans based on contact with poultry needs to be assessed properly.

In this thesis we hypothesized that the strains of human and avian origin belonging to ST95 could represent host-specific genetic variability and host specific phenotypes that can be applied to discern ST95 strains with their respective host thereby facilitating a sound risk assessment of this group of strains in future. The simple question was: are ST95 strains zoonotic or do they harbor host specific strains?

5.2. Identification of genetic markers, single nucleotide polymorphisms (SNPs) that discern ST95 strains with respect to their host origin

Before the start of the thesis our group observed that the *E. coli* from humans and avian hosts assigned to ST95 were genetically similar (in terms of virulence gene profiles, phylogenetic backgrounds, and pulsed-field gel electrophoresis profiles). Therefore, the first aim of these thesis was to identify certain genetic markers, single nucleotide polymorphism (SNP) that discern ST95 strains in a host specific manner, by analyzing the maximum common genome (MCG) and flexible or accessory gene pool to pinpoint certain candidate genes for SNP analysis. As the term “core genome” often is understood as that genome, that is biological important to enable reproduction of a particular bacterial species and discern it from another bacterial species (Sarkar and Guttman 2004), we discarded this term. Instead, we used a more descriptive but specific term, which specifically addresses those genes, that are common in regard to the genomes analyzed: The maximum common genome (MCG). By introducing this term, it is clear to the reader that based on the bioinformatic analyses, just those genes are listed, which are shared by a given number of genomes. It just is a description of the presence or absence of genes with no further functional meaning.

Molecular genetic markers are considered to be the most powerful tool for analyzing genomes and investigating genomic variation (Chris Duran 2009). Different molecular marker technologies with their relevant advantages are known such as, restriction fragment length polymorphisms (RFLP), random amplified polymorphic DNA (RAPD), DNA amplification fingerprinting (DAF), Simple sequence repeat (SSR), sequence length polymorphism (SSLP), amplified fragment length polymorphism (AFLP), cleaved amplified polymorphic sequences (CAPS) and single nucleotide polymorphism (SNP) that help in mapping molecular coordinates within a genome(Sahu et al. 2012).

RFLPs have been regarded as the most sensitive method of genotyping; however, it requires a large quantity of genomic DNA and radioactive probes (Jantusch and Hull 1996). In RAPD marker technology, a single ten nucleotide long primer having arbitrary sequence is used to amplify multiple random loci of the genomes (Williams et al. 1990). DNA amplification fingerprinting (DAF) method utilizes a single, short five nucleotide long primer to amplify many loci by PCR (Kolchinsky et al. 1993). Simple sequence repeat (SSR) markers which are

also known as microsatellite markers, make use of the variation for tandem repeats such as (CA)_n repeats that are observed between genotypes (Weber and May 1989). Simple sequence length polymorphism (SSLP) marker uses a segment of genomic DNA sequence that accommodates a simple tandem repeat distinguishing the genotypes (Bell and Ecker 1994). Another type of marker is the cleaved amplified polymorphic sequences (CAPS) that are designed according to the RFLP of PCR amplified fragments, and is used when sequence information is not known for one of the haplotypes (Konieczny and Ausubel 1993). All these marker technologies are used to detect polymorphisms to produce molecular marker maps and physical maps for species that do not have any genome sequences (Sahu et al. 2012).

Recently, large availability of assembled whole genome sequences and declining cost of sequencing has put forward another molecular marker method known as single nucleotide polymorphism (SNP) molecular marker (Davey et al. 2011, Nielsen et al. 2011, Barbazuk and Schnable 2011). Variation in the DNA sequences is used in molecular genetic marker system and its utilization in advance research is well established and accepted scientifically (Gupta PK 2001). Molecular markers based on DNA have several advantages as they are highly heritable, easily accessible and hence Single Nucleotide Polymorphisms (SNPs) based markers are predominate recently. SNPs do not change more frequently and are considered to be stable through generations and thus low mutation rates make SNPs excellent markers for understanding genomic evolution and genetic traits (Syvanen 2001).

SNP typing has been suggested as a promising approach for discriminating *E. coli* with respect to their host specific groups and for the identification of human specific *E. coli* isolates (Sheludchenko et al. 2010). SNP- based markers have been studied for the identification of markers that discriminate inter species and intra-species of obligate bacterial parasites (*Pasteuria* spp.) (Mauchline et al. 2011). Another study (Sheludchenko et al. 2011) has reported human-specific and animal-specific SNP profile, indicating that there exist host limited specificities in the SNP profile that could be used to differentiate a group of *E.coli* from the same sequence type with respect to their host. The availability of sequenced genomes has facilitated comparative genome analysis that has been used for the assessment of genetic diversity within strain types to identify DNA markers that are associated with severe disease. However, it is of utmost importance to know that due to the high diversity of *E. coli* genomes, it is simply impossible to discern specific *E. coli* genomes by SNP analysis. In contrast, SNP

analysis is envisioned to be able to discern *E. coli* genomes that are phylogenetically related into further subtypes. This is why we chose this method for the analysis of the epidemiologically important and related strains of ST95.

Comparative studies of bacterial genomes have defined the bacterial genomes into two largely independent yet intimately intertwined genomes: the “core” and “flexible” genomes (Hacker and Carniel 2001). The core genome consists of genes that are present among all strains of a bacterial species and encode proteins that are essential for the survival of the organism. Core genome components are believed to be less prone to undergo horizontal gene transfer and can be considered as the clonal backbone of the species. The flexible genome on the other hand, consists of genes that vary among strains within a species and encode proteins that may provide survival advantage by mediating adaptation to specific niches, environments or certain hosts. The flexible genome includes different virulence-associated genes, genes associated with mobile elements and resistance genes (Sarkar and Guttman 2004, Ahmed et al. 2008). The flexible genome evolves largely through horizontal genetic exchange and regularly shuffles therefore cannot be targeted to identify candidate genes that can be used as a marker to define certain group of strains. Analysis of the core genome for identification of candidate genes for DNA markers have been reported by (Forgetta et al. 2011). In our approach for the identification of a genetic marker we therefore considered the genes that belong to the MCG shared by bacterial specie.

Thus, we hypothesized that the ST95 *E. coli* strains of human and avian origin could represent host-specific genetic variability in terms of single nucleotide polymorphisms (SNPs) that differentiate the ST95 strains with respect to their hosts. Hence, we investigated genes from MCG for identifying genes that could be used for SNP analysis. With this regard, different bioinformatics approaches were put forward to identify certain candidate genes for SNP typing.

The SNPs can only be investigated in genes that are common to each and every strain, the analyses has to be focused on core genes (Kaas et al. 2012) that had to first be identified by whole genome sequence analyses. In this concern, the core genomes of publicly available 46 *E. coli* genomes including strains of ST95 were procured from NCBI and all the genes that were shared by these 46 genomes were designated “maximum common genome (MCG)” that represented the maximum number of genes shared by all the strains investigated.

At the time of the initial analysis, a total of 46 *E.coli* genomes were available in public databases. Of these 46 genomes, four genomes (APEC_O1 (NC_008563.1), S88 (NC_011742.1), IHE3034 (NC_017628.1) and UTI89 (NC_007946.1)) belonged to ST95. Once the genes from MCG were identified, a common approach as used by several groups (Ciccarelli et al. 2006, Daubin et al. 2002, Brown et al. 2001, Brochier et al. 2005, Devulder et al. 2005) was implemented that consist alignment and concatenation of orthologous core genes resulting in phylogenetic tree by using the principle of maximum likelihood (Olsen et al. 1994). Single trees for all the genes were obtained and analyzed using TREEVIEW software. We have described in our study a computational approach for screening 2,169 conserved MCG genes from *E.coli* genomes for identification of candidate genes for SNP typing. Similar approaches have been applied earlier for studying the evolutionary and functional microbial phylogenies (Segata and Huttenhower 2011). Strains differentiation based on host specific SNP marker was formulated to develop a DNA based diagnostic tool. Other groups (Sloan et al. 2008, Spigaglia et al. 2010, Wolff et al. 2009) have developed DNA marker based diagnostic tool and have relied on known genomic regions. In our study, we analyzed MCG genes for recruiting DNA based genetic marker with a view that it is more stable, less prone to recombination (Castillo-Ramirez et al. 2011) and present ubiquitously in all the strains of a species.

Since a large number of genes were identified taking into account the above computational analysis investigation of the MCG, an alternative approach was utilized based on mutation and recombination events within the 2,169 genes of MCG to reduce the number of genes to identify more specific candidate genes for SNP analysis.

E. coli strains have higher rate of mutation and recombination events that contribute the genetic variation among these strains (Wirth et al. 2006, Corander et al. 2003). However, the recombination occurs at a higher rate than the mutation rate and is considered to be a prominent force leading to the clonal divergence in *E. coli* (Guttman DS 1994). Frequent recombination in *E. coli* obscure phylogenetic signals (Holmes et al. 1999), thereby reducing the degree of differentiation between groups and strains (Wirth et al. 2006).

Feng et al. (2008) have suggested that those genes that undergo recombination are not congruent and not useful for analysis of strain relationships. Feng and their group in their study excluded regions that have undergone recombination, and considered only genes undergoing

mutational events for SNPs studies from the whole genome (Feng et al. 2008). The same concept of exclusion of regions from whole genome under recombination for SNP analysis has been used before for example for *Y. pestis* (Achtman et al. 2004, Chain et al. 2006).

Hence, genes that undergo frequent recombination were considered to be an unfit candidate for SNP analysis and therefore, in our study we focused on the differentiation of genes based on the identification of recombination events or mutations within the genes. Therefore, genes that demonstrated recombination were rejected and only genes that have undergone some mutations were selected to be candidate genes for SNP analysis.

We used the phylogenetic approach that is considered to be a powerful mechanism for detection of recombination events (Feil 2004) within the MCG genes. The phylogenetic maximum likelihood tree constructed by concatenated alignments of 2,169 genes of MCG of 46 *E. coli* genomes was used as a reference against which each single tree of all the 2169 genes was compared and the same approach has been earlier used by (Touchon et al. 2009).

The likelihood values for the trees were used as input for CONSEL (Shimodaira and Hasegawa 2001) to perform the AU test (Shimodaira 2002) to obtain p values for the trees and later the maximum chi-squared test (Smith 1992) was used to detect the recombination within the genes as it is considered to be a powerful tool for the detection of recombination within the genes. Using this approach, we could distinguish between the recombination and mutation events among the MCG genes. Therefore, genes were selected that illustrated mutation events and did not show recombination. However, these genes did not represent any genetic variability within the APEC_O1 strain and the three strains isolated from humans of ST95. As reported, consistent genetic variations that appear in members of a single group are considered as potential candidate DNA markers of diagnostic value. (Forgetta et al. 2011). Hence, these genes did not serve to be useful candidate genes for SNP analysis.

Since the earlier approaches could not point out candidate genes for SNP analysis, another alternative method was put forward to identify candidate genes for SNP typing by comparative genome analysis of 1,300 genes from the MCG of nine strains which belong to ST95. Among these nine strains, the sequences of four ST95 strains (APEC_O1, S88, IHE3034 and UTI89) were available at NCBI and the other additional five strains (two strains isolated from humans,

two strains isolated from birds and one strain of dog origin) that belong to ST95 were sequenced by our institute.

We used comparative genome analysis of MCG genes and a similar approach has been earlier reported by Vincenzo Forgetta and group (Forgetta et al. 2011). As a result, we identified twelve candidate genes that separated the nine genomes with respect to their host and were used for further SNP analysis in ST95 strains. However, these genes also failed to differentiate ST95 strains with respect to their host as no host specific genetic variation was observed. Thus, illustrating that there exist a close similarity in the sequences of the genes of strains from human and avian host of ST95 and same SNPs appeared in both the group of strains. Supporting the fact that the APEC strains belong to the same highly pathogenic clonal group as human *E. coli* strains constituting potential zoonotic risk from APEC strains as observed by (Schouleur et al. 2007).

Mora A. and group (Mora et al. 2009) revealed that ST95 was detected in strains of animal (all bird strains) and human origin and some APEC isolates may act as potential pathogens for humans and suggested no host specificity for this type of isolates. Our SNP analyses also revealed that the ST95 strains isolated from humans and avian hosts possessed same polymorphisms in their genetic makeup and were indistinguishable with respect to their host based on SNP analysis of the MCG genes. Thus, similar genetic makeup suggests that the ST95 strains isolated from avian host are closely related to the strains isolated from humans and therefore, strains of avian origin may act as a reservoir for the virulence properties of human ExPEC strains (Rodriguez-Siek et al. 2005).

The ExPEC strains of ST95 mostly belong to the human and avian host and appeared to be genetically similar (in terms of virulence gene profiles, phylogenetic backgrounds, and pulsed-field gel electrophoresis profiles). However, ExPEC strains differentiation with respect to their host is important so as to identify their origins, reservoirs and mode of transmission. Thus, we hypothesized that the ST95 strains isolated from humans and birds could represent host-specific genetic variability that can discern ST95 strains with their respective host. We recruited comparative genomic analysis approach to identify certain DNA based SNP markers to differentiate ST95 strains. Similar approach of comparative genomics has been utilized by

Vincenzo Forgetta and group for the discovery of diagnostic DNA-based targets that are species specific or associated with multiple severe disease associated strains (Forgetta et al. 2011).

E. coli are diverse organisms, a clear definition is therefore required to group *E. coli* according to their host within a sequence type as same sequence type is shared by the strains from different host. DNA based diagnostic tool can thereby differentiate strains according to their host of origin and facilitate a sound risk assessment of these ST95 strains in future. Host specific SNP could help in understanding the role of genetic variability and specific polymorphism that facilitates bacterial adaptation towards host environment and survival strategies implicated by bacteria thereby leading to infection.

However, our results falsifies our hypothesis and do not indicate any host specific SNP among ST95 strains. Therefore, it can be concluded that the ST95 strains from different host may have similar genetic makeup and may be the descendent of common ancestors. Similar genetic and ancestral credential may support the fact that the ST95 strains of avian origin can have zoonotic potential. The avian ExPEC strains have also been suggested to have zoonotic potential by other investigators (Schouleur et al. 2006, Schouleur et al. 2007, Ewers et al. 2007, Johnson et al. 2008d, Ewers et al. 2009).

As the MCG did not demonstrate any host specific SNP, we also assessed the gene that has been used earlier to discern ExPEC groups. The *fimH* gene was chosen because of the following reasons: (i) the typeability of this gene is easy as it is present in both commensal and pathogenic strains of *E. coli* (ii) *fimH* gene has been studied among ExPEC strains, in particular of human urinary pathogenic *E. coli*, and sequence variations in *fimH* gene have been determined as pathoadaptive mutations conferring differing adhesion capabilities according to various host (Tartof et al. 2007).

Previous studies showed, *fimH* SNP analysis as a possible screening tool to type UPEC (Tartof et al. 2007) thereby making it a simple, sequence-based screening method for differentiating strains of different host. The *fimH* gene genotyping along with internal fragments of *fumC* gene that is used in MLST have been studied to have higher clonal discrimination power (Weissman et al. 2012). However, our findings suggest that the *fimH* SNP typing could not be used to differentiate strains of ST95 with respect to their human and avian hosts, as the major allelic

groups that were identified, contained a mixed group of strains from different hosts in same allelic group. No host specific pathoadaptive mutations were observed. Therefore, *fimH* SNP typing proved to be an inappropriate tool for distinguishing strains of ST95.

Fourteen different *fimH* alleles were obtained as a result of *fimH* gene SNP analysis. The major allelic groups possessed ST95 strains of different host together. We were interested to investigate the phenotypic characters of some strains from predominant allelic variants with respect to the *fimH* gene as amino acid variation within the adhesin molecule that may contribute to the variation in receptor binding ability and receptor recognition and are considered as allelic variation or pathoadaptive mutations. Acquisition of mutations within *fimH* adhesin gene reflects the adaptation strategies of *E. coli* for its existence in alternative habitats during infection and colonization within a host (Weissman et al. 2006). It has been known that the phenotypic variants of FimH are predominantly the product of SNPs in *fimH* (Hommais et al. 2003) and some of the *fimH* gene sequence variations are responsible for the increased adhesiveness and binding of *E. coli* strains to monomannose residues, thus playing a role in extraintestinal infections (Bauchart et al. 2010). Also the genetic variation in the *fimH* gene of type 1 fimbriae mediates changes in the tropism of *E. coli*, shifting it towards a virulent phenotype (Sokurenko et al. 2004). Thus, FimH adhesin has turned out to be an important attribute mediating colonizing of *E. coli* in different niche and can prove to be a useful model to study single gene adaptations as it is known that different structural variants of FimH vary in the strength of their binding to uroepithelial cells (Sokurenko et al. 1997, Sokurenko et al. 1995).

In conclusion, our corroborative data did not give any evidence for specific genetic difference, able to discern ST95 ExPEC strains from human and avian hosts. Therefore, in a second step we performed functional assays of these same strains. This rational was based on the hypothesis that if the genetic makeup of the strains is not different, gene expression in functional assays might reveal host-specific features.

5.3. Identification of certain phenotypic characteristics in ST95 strains that could segregate this group of strains according to their human/avian host respectively

In this regard, different phenotypic characteristics like adhesion and invasion capabilities, biofilm forming capabilities, expression of curli and cellulose via long term biofilm colony

formation, haemagglutination assay and resistivity of ST95 strains towards chicken and human serum were accessed, so as to investigate if the avian strains demonstrate different phenotypic behavior from that of the human strains of ST95.

Adherence to HEP-2 cells is considered to be the most useful phenotypic assay to diagnose diarrheagenic *E. coli* (Donnenberg and Nataro 1995) and remains a “gold standard” for the diagnosis of EAEC and diffusely adherent *E. coli* (DAEC) (Nataro and Kaper 1998). Adherent invasive *E. coli* (AIEC) identification is also based on their ability to adhere and invade intestinal epithelial cells (Darfeuille-Michaud et al. 2004, Martinez-Medina et al. 2011). Thus, adherence and the invasion assays were reported as phenotypic assays to identify different pathotypes of *E. coli*. Human ExPEC particularly, UPEC strains adhere human bladder epithelial cells (Eto et al. 2007) and ascend towards kidney (Chassin et al. 2011). Hence, we used the adhesion and the invasion assay to investigate if the ST95 strains from humans and avian host demonstrate differences in their adherence and invasion pattern.

Biofilm formation is an important feature in the pathogenesis of human ExPEC strains particularly for the UPEC strains promoting their persistence in the bladder epithelial cells (Soto et al. 2007). Several groups have reported biofilm formation by UPEC strains (Soto et al. 2006, Ulett et al. 2007, Ong et al. 2008). Recently biofilm formation has been studied as a novel phenotypic character of AIEC strains (Margarita et al. 2009). Thus, we studied the biofilm formation capabilities of ST95 strains under different media conditions.

Curli fibres are involved in biofilm formation and mediate adhesion and invasion (Barnhart and Chapman 2006, Saldana et al. 2009). Curli production has been studied earlier in human and avian ExPEC strains (Provence and Curtiss 1992, Bian et al. 2000). However the expression of curli is cryptic at different temperature (Beloin et al. 2008). Hence, we studied the expression of curli and cellulose under different temperature and media conditions in ST95 strains of human and avian origin to investigate if these strains show differences in curli and cellulose expression under different conditions.

We also assessed the expression of type 1 fimbriae among the strains of ST95 as it is an important virulence factor that is involved in the initial stage of biofilm formation and also promote bacterial adhesion to the host epithelial cells (Pruss et al. 2006). Biofilm forming *E. coli* strains show more expression of type 1 fimbriae than the non biofilm forming strains (Soto

et al. 2007). Since we investigated the biofilm forming capacity of ST95 strains we also checked the expression of type 1 fimbriae via haemagglutination assay in presence of mannose.

Serum resistance is known to contribute to the pathogenicity of APEC strains (Mellata et al. 2003). Ewers. C. and group have shown in their study that nearly 80-90% of outbreak and non outbreak avian strains represent serum resistant phenotype (Ewers et al. 2009). Therefore, we investigated the serum resistant phenotype among the ST95 strains of human and avian origin to determine if these strains can be differentiated based on their serum resistant phenotypic character.

5.3.a. Adhesion and invasion capabilities of human and avian strains of ST95

The adhesion and the invasion capabilities of thirteen strains isolated from humans (six NMEC, six UPEC strains and one commensal strain) and eight strains isolated from birds (six APEC strains and two commensal strains) out of 116 ST95 strains were investigated in HEK293-T (Human embryonic kidney) cell line and MDCK-1 (Madin-Darby Canine Kidney-1) cell line. The strain having similar virulence gene pattern and clonal nature as observed by their ST and similar macro-restriction pattern of the genome (Pulsed-field gel electrophoresis) that had been done before the start of this thesis by our groups were used for the infection experiment. Human and dog kidney cell line was used for infection experiments considering the background knowledge that UPEC strains adhere to bladder epithelial cells and ascend towards kidney (Chassin et al. 2011); hence we investigated if the ST95 strains of humans adhere to the kidney cells and if the same phenotype is also demonstrated by the ST95 strains isolated from birds. The dog cell line was also used so as to investigate if the ST95 strains can adhere to and invade kidney cells from host other than humans and birds as recently two ST95 strains were isolated from dog.

UPEC have been extensively studied with respect to their adhesion and invasion mechanisms towards bladder epithelial cells and are known to adhere and invade kidney cells by ascending through the urinary tract causing pyelonephritis (Chassin et al. 2011). Stehling EG (Stehling EG 2003) have studied APEC adhesion and invasion in Hep 2 cells and the APEC strains are also known to invade chicken embryo fibroblast (Wang et al. 2011). By observing the adhesion and invasion patterns of ST95 strains isolated from human and avian hosts in our study we could

conclude that the strains from birds can adhere and invade human cell line and were more invasive towards human kidney cells as compared to the strains isolated from humans. However, no significant difference was observed in the adhesion pattern of strains from birds when tested against strains isolated from humans.

Therefore, it can be concluded that the ST95 strains isolated from birds were indistinguishable from strains isolated from humans phenotypically in terms of their adherence and invasive properties. As mentioned in previous studies, certain APEC strains were found to be highly similar to human ExPEC (Ewers et al. 2009) and poultry was considered to be a reservoir of ExPEC with zoonotic potential (Belanger et al. 2011). The higher invasion phenotypes of strains of avian origin towards human cell line also support their potential zoonotic nature making these strains more successful pathogens. The adherence characteristic of strains isolated from avian and human hosts towards MDCK-1 cells indicates the fact that the ExPEC strains of ST95 can cross the host species barrier and colonize and infect host other than human and avian group and therefore broadening their host spectrum.

5.3.b. Human and avian strains of ST95 represent similar biofilm forming capabilities

Biofilm was defined by Costerton (Costerton 1999) as a structured community of bacterial cells that is enclosed in a self-produced polymeric matrix adhering an inert or living surface. Biofilm formation is an important feature of UPEC strains and has been studied extensively for their ability to form biofilm. Biofilm promotes persistence of UPEC strains in the urinary tract protecting bacteria from the cleansing out effect of hydrodynamic forces and escaping the killing activity of host defence mechanisms and protection against antibiotics (Hanna et al. 2003). In total, 116 ST95 strains (used in all further experiments) including 61 strains isolated from birds (58 APEC strains and 3 commensal strains) and 55 strains isolated from humans (24 UPEC, 15 NMEC, 4 SePEC and 15 commensal strains) available at our institute were investigated for biofilm formation capabilities under different media (LB media and M63 media) conditions. Strains isolated from humans more specifically strains belonging to UPEC pathotype are known to form biofilm (Mulvey et al. 2001, Eto et al. 2007). We investigated, if the biofilm forming phenotype could be implemented for the differentiation of strains isolated from birds and humans in accordance with their host. Our results suggest that there was no significant difference in the biofilm formation capabilities of strains of birds when compared to

that of the strains of humans. Thus, it can be concluded that the strains isolated from birds and humans demonstrated similar phenotypes in terms of their biofilm formation capabilities.

We also observed that both the groups of strains produced more biofilm in nutrient depleted media like M63 as compared to the nutrient rich media like LB media and the same tendency of strains producing more biofilm in nutrient depleted media was observed in previous studies by Skyberg (Skyberg et al. 2007). Our results revealed varying abilities of strains of human and avian origin to form biofilm in vitro. As documented in earlier studies, the growth medium plays a significant role in the biofilm formation capabilities (Sheikh et al. 2001, Reisner et al. 2006). The biofilm formation in nutrient depleted media was higher than nutrient rich media suggesting that the biofilm formation capacity increases in media with low nutrient and lower osmolarity. APEC strains also form biofilm and it has been studied that there occurs the transfer of genetic elements in the biofilm (Davey and O'Toole G 2000). This transfer of genetic material can also lead to the exchange of antibiotic resistant and virulence plasmids. Previous studies with animal experiments suggested that the APEC plasmids were capable of contributing to the urovirulence in mammalian hosts, further suggesting that the APEC plasmids could be reservoirs of virulence genes for UTI causing *E. coli* in human (Skyberg et al. 2006).

Thus, the biofilm formation capacity of bacteria turned out to be media dependent and the ST95 strains isolated from humans and birds are indistinguishable in terms of their biofilm formation in the media tested. Hence, this phenotype of biofilm formation does not serve as a candidate phenotype to discern strains of ST95. The exchange of antibiotic genes and virulence genes among APEC strains during the course of biofilm formation and their dissemination into the environment could lead to the emergence of more successful pathogens that could even develop zoonotic potential causing infection in humans.

5.3.c. Curli and cellulose expression (Long term biofilm formation)

The expression of curli and cellulose was also accessed via long term biofilm formation by the ST95 strains of human and avian origin to investigate if the differential expression of curli and cellulose could be utilized as a phenotypic marker to discern strains of ST95.

Curli fimbriae have been demonstrated to play an important role in adhesion to the host cells and abiotic surfaces and are also involved in biofilm formation by many Enterobacteria (Cookson et al. 2002, Brombacher et al. 2006). Curli have been studied to attach to the proteins of extracellular matrix such as plasminogen, fibronectin and laminin thereby promoting adhesion of the bacteria to different human cells (Olsen et al. 1989, Ben Nasr et al. 1996). Expression of curli is also linked to cellulose biosynthesis, leading to the production of extracellular matrix resulting in tight, cell to cell and cell to surface interactions and producing a colony morphotype called rdar (red, dry and rough) (Brombacher et al. 2006). Earlier studies have revealed that the cellulose production in bacteria is mainly associated with the ability to form a rigid biofilm however, these phenomenon vary from strain to strain and different serovars and is highly influenced by environmental conditions (Beloin et al. 2008).

Colony morphotypes were studied to determine the expression of curli and cellulose (Zogaj et al. 2001, Romling 2005). The red, dry and rough morphotype abbreviated as Rdar was demonstrated when cellulose and curli are expressed together. The expression of only cellulose by bacteria was represented by pink, dry and rough (pdar) colony morphotype whereas, expression of only curli was represented by brown, dry and rough (bdar) morphotype. When neither curli nor cellulose was expressed the colony appeared to be smooth and white and this phenotypic morphotype is named as saw. Different colony morphotypes were represented by ST95 strains on agar plates supplemented with Congo red dye during the expression of curli and cellulose.

Bian et al in their studies have shown that during urinary tract infection curli appears in large aggregates in uropathogenic *E. coli* and promote bacterial adherence within urinary tract (Bian et al. 2000). Another study has shown that curli promotes adherence of avian pathogenic *E. coli* isolates to avian intestinal cells and internalization and persistence in the caecum of chickens (Saldana et al. 2009). Cellulose along with curli, form special structures that enable biofilm development through bacterial interactions and adhesion to biotic and abiotic surfaces, providing protection against environmental stress thus, characterizing curli as a virulence attribute (Bian et al. 2000, Saldana et al. 2009).

Curli synthesis in clinical isolates of *E. coli* is under complex regulation promoting curli production notably at 37°C and/or 28°C depending on the isolates, although curli expression is

cryptic in most *E. coli* (Beloin et al. 2008). In our study, we found that both the group of strains isolated from humans and avian hosts illustrated expression of curli and cellulose. However, the expression of curli and cellulose together was observed only at lower temperature of 28°C and only on plates without salt, thus suggesting that the expression of curli and cellulose together in ST95 strains, occurred at lower temperature as observed by Gualdi and group (Gualdi et al. 2008) and at lower osmotic conditions.

We also observed only curli expression by strains of humans and avian origin at 28°C on plates with salt and without salt, but at 37°C the expression of curli was observed only on plates without salt and not on the plates with salt incubated at 37°C. This corroborates with the previous studies that suggest that most of the pathogenic *E. coli* strains do not express curli at higher temperature of 37°C but at temperatures below 30°C and under low nutrients and low media osmolarity, during the stationary growth phase (Olsen et al. 1993, Uhlich et al. 2001).

Though, most of the *E. coli* strains carry genes for curli expression, only a subset of them can transcribe the gene that is often suppressed when the *E. coli* is grown at 37°C (Olsen et al. 1993, Uhlich et al. 2001). Saldana and group (Saldana et al. 2009) have described in their studies that heterologous serotypes of pathogenic *E. coli* were able to assemble curli fibers while growing on salt-less media plates at 37°C supporting our results of curli expression at 37°C on plates without salt and not on plates with salt. Though it has been documented that the curli and cellulose expression at 28°C is predominant in UTI isolates (Bokranz et al. 2005) the avian strains of ST95 also illustrated similar phenotypic expression as that of human strains.

As the ST95 strains isolated from birds demonstrated similar expression patterns as that of the strains of humans in terms of their curli and cellulose expression, under different sets of conditions tested, it can be concluded that both the groups of strains are indistinguishable and different morphotypes observed during curli and cellulose expression did not prove to be a distinctive feature separating ST95 strains.

5.3.d. Haemagglutination assay among ST95 strains

Characterization of bacterial adhesins is done according to the agglutination patterns as a result of bacteria binding to the erythrocytes from various species. Bacteria that agglutinate erythrocytes and do not show this phenomenon in the presence of D-mannose were defined as

carriers of type 1 pili (Hagberg et al. 1981). Type 1 fimbriae are a crucial factor for the virulence of uropathogenic *E. coli* mediating adhesion to epithelial cells and colonization of the tissues and for invasion of the uroepithelium (Muller et al. 2009) and thus, assigning it as an important virulence attribute. Haemagglutination assay was performed with chicken erythrocytes in presence and absence of D-mannose for the strains of ST95 in order to investigate if these strains show haemagglutination defined as Mannose Sensitive Haemagglutination (MSHA) or Mannose Resistant Haemagglutination (MRHA).

Both the groups of strains isolated from humans and birds did not represent mannose resistant haemagglutination (MRHA). Earlier studies have also demonstrated that the APEC strains causing colibacillosis employ type 1 fimbriae for adherence and tracheal colonization (Pourbakhsh et al. 1997, Edelman et al. 2003) and this supports our result as all the avian strains that demonstrated agglutination were pathogenic strains that were associated with colibacillosis. Therefore, the strains of human and avian origin demonstrated similar agglutination pattern and there was no difference in the haemagglutination properties. Thus, it can be concluded that the mannose specific haemagglutination could not serve to discern ST95 strains.

5.3.e. Serum resistance of ST95 strains in chicken and human serum

The serum resistant phenotype of ST95 strains in chicken and human serum was also investigated. Serum has a lethal effect on Gram negative bacteria and the bactericidal activity of serum is due to the complement system of serum that plays an important role in the host defense mechanism against foreign invasion. Bacterial resistance to the bactericidal effect of serum was shown by many pathogenic strains and particularly by invasive strains and those causing UTIs (Burke et al. 1990). Human ExPEC and avian *E. coli* strains possess several virulence factors that facilitate bacterial resistance against the complement system of the host (Wooley et al. 1993).

Resistance of *E. coli* to human serum is regarded as an important virulence attribute and is mediated by O-antigen polysaccharide side chains (Stawski et al. 1990), capsular polysaccharides (Leying et al. 1990) and outer membrane lipoproteins (Siegfried et al. 1995). Serum resistance was also observed in APEC strains as an important virulence determinant that

plays a major role in the pathogenesis of avian colibacillosis (Mellata et al. 2003) and was used to make distinction between virulent strains and non virulent strains (Nolan et al. 2003).

In the present study we investigated serum resistance of ST95 strains isolated from humans and birds (strains from different pathotypes) in chicken and human serum. To our observation, we found that most of the strains of human and avian origin demonstrated resistant phenotype towards the bactericidal effect of chicken serum. Ewers C. and group have also shown in their study that nearly 80-90% of outbreak and non outbreak avian strains represented serum resistant phenotype (Ewers et al. 2009).

The higher percentage of strains representing resistant phenotype to the serum complement system can be explained by the fact that these strains possess the ability to adapt themselves rapidly to the stress that is encountered during the incubation of strains in serum. Ganwu Li and their group (Li et al. 2011) have shown in their study that the strains adapt themselves in response to serum and several genes involved in stress resistance like acid, oxidative stress resistance genes were upregulated to counter act oxidative stress during growth in chicken serum. Li G. and group (Li et al. 2012) have also shown that the avian pathogenic *E. coli* strain and the neonatal meningitis *E. coli* used similar mechanisms to overcome the antibacterial effects of sera, when grown in human and avian serum and notably, several proteins involved in synthesis of sulfur-containing amino acids and fatty acids, nucleic acids, were expressed differentially in response to the sera from different hosts. And this could be the possible reason of differential resistant phenotype of ST95 strains in chicken and human serum.

Therefore, the ST95 strains isolated from humans and birds were indistinguishable in terms of their serum resistant phenotype toward chicken and human serum and therefore, the serum complement resistant phenotype does not serve as a promising phenotypic characteristic to discern strains of ST95 with respect to their host and this corroborates with the earlier studies of (Ewers et al. 2009) revealing the fact that the serum resistance was not a useful marker to distinguish avian *E coli* from different source groups. And, as the strains of avian origin are competent to utilize the same strategies as that of the strains isolated from humans to overcome the bactericidal effect of serum from different hosts, it aids to the possible zoonotic nature of strains isolated from birds thereby potentially making avian strains capable of causing infections in other hosts.

6. Conclusion

In conclusion, we were not able to discern any significant features of the 116 strains we analyzed in this promotion thesis, neither by genomic based in-silico nor by functional biological analysis, distinguishing avian from human *E. coli* strains of ExPEC ST95. Due to these data, ExPEC strains of ST95 have to be defined as zoonotic agents. To our knowledge, such a strong evidence for any ExPEC to be zoonotic has not been published before. As these strains investigated belong to 43 different serotypes, we consider the ST to be of more significance to judge on the biological features than the conventional typing methods.

We thereby largely extended previous studies reporting on similar genetic makeup and indistinguishable phenotypic characters and commonality in serogroups, phylogenetic background, virulence genes profile and the ability to cause disease in certain animal models by human ExPEC and APEC isolates (Wirth et al. 2006, Johnson et al. 2007, Moulin-Schouleur M 2007, Ewers et al. 2009) explains that the avian strains of ST95 are indistinguishable from human ExPEC strains and indicates the fact, that ExPEC strains of ST95 possess zoonotic potential. As in particular chickens are held in large populations and avian pathogenic *E. coli* is a prominent cause of disease in these animals, we consider in particular chickens as a relevant risk of human infections via food or direct contact. Furthermore, these APEC, UPEC and NMEC have common ancestors.

Recently, two strains of ST95 were observed in dogs which belonged to the same phylogenetic group B₂ and harbored similar sets of virulence genes as that of the other human and avian strains of ST95 (Ewers, Wieler et al., unpublished data). UPEC strains isolated from dogs have been observed earlier to be phylogenetically related to that of the humans (Johnson et al. 2001b, Johnson et al. 2008b). Such strains are thought to cross the host specificity barrier and colonize and infect humans and other companion animals such as cats as observed in the strains of other STs as ST131 (Johnson et al. 2009). Earlier reports have shown that a UPEC strain causing UTI in dog was also found to colonize humans in the same residence (Johnson et al. 2008a) illustrating animal-to-human transmission of ExPEC. Therefore, this is indicating the fact that companion animals could be potential reservoirs for the transmission of UPEC to humans.

7. Zusammenfassung

Im Verlauf dieser Promotionsarbeit war es uns nicht möglich, signifikante Eigenschaften der 116 analysierten Stämme durch Genom basierte *in silico* sowie funktionelle biologische Analysen zu ermitteln, die eine Unterscheidung von aviären zu humanen *E.coli* Isolaten von ExPEC ST95 ermöglichen. Auf Grund dieser Daten sollten ExPEC Stämme des ST95 als Zoonoseerreger definiert werden. Nach unserem Wissen wurde ein solch starker Beweis für das zoonotische Potential von ExPEC bisher noch nicht veröffentlicht. Da die untersuchten Stämme 43 verschiedenen Serotypen zugeordnet sind, betrachten wir den Sequenztyp (ST) im Vergleich zu den herkömmlich verwendeten Typisierungsmethoden als genauer um biologische Eigenschaften beurteilen zu können.

Dabei haben wir größtenteils auf Ergebnisse früherer Studien aufgebaut, die den ähnlichen genetischen Aufbau und die nicht zu unterscheidenden phänotypischen Eigenschaften und Gemeinsamkeiten in Serogruppen, den phylogenetischen Hintergrund, das Virulenzgen-Profil und die Fähigkeit von humanen ExPEC und APEC Krankheiten in bestimmten Tiermodellen zu verursachen, aufzeigen (Wirth et al. 2006, Johnson et al. 2007, Moulin-Schouleur M 2007, Ewers et al. 2009) und somit erklären, dass aviäre Stämme des ST95 nicht von humanen ExPEC zu unterscheiden sind und ExPEC des ST95 somit ein zoonotisches Potential besitzen.

Da vor allem Hühner in großen Populationen gehalten werden und aviäre pathogene *E.coli* eine bedeutende Ursache für Krankheiten in diesen Tieren sind, betrachten wir besonders Hühner als relevanten Risikofaktor für die Infektion von Menschen durch Lebensmittel oder direkten Kontakt. Weiterhin haben APEC, UPEC und NMEC gemeinsame Vorgänger.

Kürzlich wurden zwei Stämme des ST95 in Hunden entdeckt, die zur selben phylogenetischen Gruppe B2 gehören und beide die gleiche Zusammenstellung an Virulenzgenen aufweisen wie andere humane und aviäre Isolate des ST95 (Ewers, Wieler et al., unpublished data). UPEC Stämme, die von Hunden isoliert wurden, konnten schon früher in phylogenetischen Zusammenhang mit humanen Isolaten gebracht werden (Johnson et al. 2001b, Johnson et al. 2008b). Von solchen Stämmen wird vermutet, dass sie die wirtsspezifische Grenze überqueren und so Menschen und andere Haustiere wie z.B. Katzen kolonisieren und infizieren können, wie es schon bei anderen STs wie ST131 beobachtet werden konnte (Johnson et al. 2009). Frühere Berichte haben gezeigt, dass UPEC Stämme die in Hunden UTI verursachen auch die

mit ihnen zusammenlebenden Menschen kolonisieren können (Johnson et al. 2008a), was die Tier-zu-Mensch Übertragung von ExPEC unterstreicht. Diese Fakten weisen darauf hin, dass Haustiere ein potentielles Reservoir für die Übertragung von UPEC auf den Menschen sind.

8. References

- ABDUL-TEHRANI H, HUDSON AJ, CHANG YS, TIMMS AR, HAWKINS C, WILLIAMS JM, HARRISON PM, GUEST JR AND ANDREWS SC. 1999. Ferritin mutants of *Escherichia coli* are iron deficient and growth impaired, and *fur* mutants are iron deficient. *J Bacteriol* 181: 1415-1428.
- ACHTMAN M ET AL. 2004. Microevolution and history of the plague bacillus, *Yersinia pestis*. *Proc Natl Acad Sci U S A* 101: 17837-17842.
- AGUERO ME, ARON L, DELUCA AG, TIMMIS KN AND CABELLO FC. 1984. A plasmid-encoded outer membrane protein, TraT, enhances resistance of *Escherichia coli* to phagocytosis. *Infect Immun* 46: 740-746.
- AHMED N, DOBRINDT U, HACKER J AND HASNAIN SE. 2008. Genomic fluidity and pathogenic bacteria: applications in diagnostics, epidemiology and intervention. *Nat Rev Microbiol* 6: 387-394.
- ALLEN PM, ROBERTS I, BOULNOIS GJ, SAUNDERS JR AND HART CA. 1987. Contribution of capsular polysaccharide and surface properties to virulence of *Escherichia coli* K1. *Infect Immun* 55: 2662-2668.
- ANANIAS M AND YANO T. 2008. Serogroups and virulence genotypes of *Escherichia coli* isolated from patients with sepsis. *Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas / Sociedade Brasileira de Biofisica [et al]* 41: 877-883.
- ANDREOLI CCJC, ROBERT GRIGGS, JOSEPH LOSCALZO 1997. *Cecil Essentials of Medicine*. Philadelphia: WB Saunders 4th edition.
- ANDREWS SC, ROBINSON AK AND RODRIGUEZ-QUINONES F. 2003. Bacterial iron homeostasis. *FEMS microbiology reviews* 27: 215-237.
- ANNANE D, BELLISSANT E AND CAVAILLON JM. 2005. Septic shock. *Lancet* 365: 63-78.
- ANTAO EM, WIELER LH AND EWERS C. 2009. Adhesive threads of extraintestinal pathogenic *Escherichia coli*. *Gut Pathog* 1: 22.
- AUSTIN PF, CASALE AJ, CAIN MP, RINK RC AND WEINTRAUB SJ. 2003. Lipopolysaccharide and inflammatory cytokines cause an inducible nitric oxide synthase-dependent bladder smooth muscle fibrotic response. *The Journal of urology* 170: 645-648.
- BABAI R, BLUM-OEHLER G, STERN BE, HACKER J AND RON EZ. 1997. Virulence patterns from septicemic *Escherichia coli* O78 strains. *FEMS Microbiol Lett* 149: 99-105.
- BACKHED F, ALSÉN B, ROCHE N, ANGSTROM J, VON EULER A, BREIMER ME, WESTERLUND-WIKSTROM B, TENEBERG S AND RICHTER-DAHLFORS A. 2002. Identification of target tissue glycosphingolipid receptors for uropathogenic, F1C-fimbriated *Escherichia coli* and its role in mucosal inflammation. *J Biol Chem* 277: 18198-18205.
- BACKHED F, NORMARK S, SCHWEDA EK, OSCARSON S AND RICHTER-DAHLFORS A. 2003. Structural requirements for TLR4-mediated LPS signalling: a biological role for LPS modifications. *Microbes Infect* 5: 1057-1063.
- BARBAZUK WB AND SCHNABLE PS. 2011. SNP discovery by transcriptome pyrosequencing. *Methods Mol Biol* 729: 225-246.
- BARNHART MM AND CHAPMAN MR. 2006. Curli biogenesis and function. *Annual review of microbiology* 60: 131-147.

- BAUCHART P, GERMON P, BREE A, OSWALD E, HACKER J AND DOBRINDT U. 2010. Pathogenomic comparison of human extraintestinal and avian pathogenic *Escherichia coli*--search for factors involved in host specificity or zoonotic potential. *Microb Pathog* 49: 105-115.
- BELANGER L, GARENAUX A, HAREL J, BOULIANNE M, NADEAU E AND DOZOIS CM. 2011. *Escherichia coli* from animal reservoirs as a potential source of human extraintestinal pathogenic *E. coli*. *FEMS immunology and medical microbiology* 62: 1-10.
- BELL CJ AND ECKER JR. 1994. Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. *Genomics* 19: 137-144.
- BELOIN C, ROUX A AND GHIGO JM. 2008. *Escherichia coli* biofilms. *Current topics in microbiology and immunology* 322: 249-289.
- BEN NASR A, OLSEN A, SJOBRING U, MULLER-ESTERL W AND BJORCK L. 1996. Assembly of human contact phase proteins and release of bradykinin at the surface of curli-expressing *Escherichia coli*. *Mol Microbiol* 20: 927-935.
- BERGER CN, BILLKER O, MEYER TF, SERVIN AL AND KANSAU I. 2004. Differential recognition of members of the carcinoembryonic antigen family by Afa/Dr adhesins of diffusely adhering *Escherichia coli* (Afa/Dr DAEC). *Mol Microbiol* 52: 963-983.
- BIAN Z, BRAUNER A, LI Y AND NORMARK S. 2000. Expression of and cytokine activation by *Escherichia coli* curli fibers in human sepsis. *The Journal of infectious diseases* 181: 602-612.
- BIDET P, BONACORSI S, CLERMONT O, DE MONTILLE C, BRAHIMI N AND BINGEN E. 2005. Multiple insertional events, restricted by the genetic background, have led to acquisition of pathogenicity island III96-like domains among *Escherichia coli* strains of different clinical origins. *Infect Immun* 73: 4081-4087.
- BIDET P, METAIS A, MAHJOUB-MESSAI F, DURAND L, DEHEM M, AUJARD Y, BINGEN E, NASSIF X AND BONACORSI S. 2007. Detection and identification by PCR of a highly virulent phylogenetic subgroup among extraintestinal pathogenic *Escherichia coli* B2 strains. *Appl Environ Microbiol* 73: 2373-2377.
- BINGEN E, PICARD B, BRAHIMI N, MATHY S, DESJARDINS P, ELION J AND DENAMUR E. 1998. Phylogenetic analysis of *Escherichia coli* strains causing neonatal meningitis suggests horizontal gene transfer from a predominant pool of highly virulent B2 group strains. *The Journal of infectious diseases* 177: 642-650.
- BISTER B, BISCHOFF D, NICHOLSON GJ, VALDEBENITO M, SCHNEIDER K, WINKELMANN G, HANTKE K AND SUSSMUTH RD. 2004. The structure of salmochelins: C-glucosylated enterobactins of *Salmonella enterica*. *Biometals* 17: 471-481.
- BLUM G, FALBO V, CAPRIOLI A AND HACKER J. 1995. Gene clusters encoding the cytotoxic necrotizing factor type 1, Prs-fimbriae and alpha-hemolysin form the pathogenicity island II of the uropathogenic *Escherichia coli* strain J96. *FEMS Microbiol Lett* 126: 189-195.
- BLYTON MD, BANKS SC, PEAKALL R AND GORDON DM. 2013. High temporal variability in commensal *Escherichia coli* strain communities of a herbivorous marsupial. *Environmental microbiology* 15: 2162-2172.
- BOGYIOVA E, KMETOVA M, BIROS E AND SIEGFRIED L. 2002. Detection of pap-, sfa- and afa-specific DNA sequences in *Escherichia coli* strains isolated from extraintestinal material. *Folia microbiologica* 47: 723-726.

- BOKRANZ W, WANG X, TSCHAPE H AND ROMLING U. 2005. Expression of cellulose and curli fimbriae by *Escherichia coli* isolated from the gastrointestinal tract. *J Med Microbiol* 54: 1171-1182.
- BONACORSI S, CLERMONT O, HOUDOUIN V, CORDEVANT C, BRAHIMI N, MARECAT A, TINSLEY C, NASSIF X, LANGE M AND BINGEN E. 2003. Molecular analysis and experimental virulence of French and North American *Escherichia coli* neonatal meningitis isolates: identification of a new virulent clone. *The Journal of infectious diseases* 187: 1895-1906.
- BONACORSI S, LEFEVRE S, CLERMONT O, HOUDOUIN V, BOURRILLON A, LOIRAT C, AUJARD Y AND BINGEN E. 2005. *Escherichia coli* strains causing urinary tract infection in uncircumcised infants resemble urosepsis-like adult strains. *The Journal of urology* 173: 195-197; discussion 197.
- BOUCKAERT J ET AL. 2005. Receptor binding studies disclose a novel class of high-affinity inhibitors of the *Escherichia coli* FimH adhesin. *Mol Microbiol* 55: 441-455.
- BRADY H. 1999. *Endotoxin in Health and Disease*. Marcel Dekker, Inc, New York edition 1.
- BROCHIER C, FORTERRE P AND GRIBALDO S. 2005. An emerging phylogenetic core of Archaea: phylogenies of transcription and translation machineries converge following addition of new genome sequences. *BMC evolutionary biology* 5: 36.
- BROMBACHER E, BARATTO A, DOREL C AND LANDINI P. 2006. Gene expression regulation by the Curli activator CsgD protein: modulation of cellulose biosynthesis and control of negative determinants for microbial adhesion. *J Bacteriol* 188: 2027-2037.
- BROWN JR, DOUADY CJ, ITALIA MJ, MARSHALL WE AND STANHOPE MJ. 2001. Universal trees based on large combined protein sequence data sets. *Nature genetics* 28: 281-285.
- BRZUSKIEWICZ E, GOTTSCHALK G, RON E, HACKER J AND DOBRINDT U. 2009. Adaptation of Pathogenic *E. coli* to Various Niches: Genome Flexibility is the Key. *Genome dynamics* 6: 110-125.
- BUCKLES EL, WANG X, LANE MC, LOCKATELL CV, JOHNSON DE, RASKO DA, MOBLEY HL AND DONNENBERG MS. 2009. Role of the K2 capsule in *Escherichia coli* urinary tract infection and serum resistance. *The Journal of infectious diseases* 199: 1689-1697.
- BULLEN JJ, ROGERS HJ, SPALDING PB AND WARD CG. 2005. Iron and infection: the heart of the matter. *FEMS immunology and medical microbiology* 43: 325-330.
- BURKE DA, CLAYDEN SA AND AXON AT. 1990. Serum bactericidal resistance of faecal *Escherichia coli* and bactericidal competence of serum from patients with ulcerative colitis. *Journal of clinical pathology* 43: 316-319.
- CAM LL. 1990. Maximum Likelihood: an introduction. *International Statistical Institute* 58,2: 153-171.
- CAPRIOLI A, FALBO V, RUGGERI FM, BALDASSARRI L, BISICCHIA R, IPPOLITO G, ROMOLI E AND DONELLI G. 1987. Cytotoxic necrotizing factor production by hemolytic strains of *Escherichia coli* causing extraintestinal infections. *Journal of clinical microbiology* 25: 146-149.
- CARNIEL E, GUILVOUT I AND PRENTICE M. 1996. Characterization of a large chromosomal "high-pathogenicity island" in biotype 1B *Yersinia enterocolitica*. *J Bacteriol* 178: 6743-6751.
- CASTILLO-RAMIREZ S, HARRIS SR, HOLDEN MT, HE M, PARKHILL J, BENTLEY SD AND FEIL EJ. 2011. The impact of recombination on dN/dS within recently emerged bacterial clones. *PLoS pathogens* 7: e1002129.

- CAVALIERI SJ, BOHACH GA AND SNYDER IS. 1984. *Escherichia coli* alpha-hemolysin: characteristics and probable role in pathogenicity. *Microbiological reviews* 48: 326-343.
- CHAIN PS, HU P, MALFATTI SA, RADNEDGE L, LARIMER F, VERGEZ LM, WORSHAM P, CHU MC AND ANDERSEN GL. 2006. Complete genome sequence of *Yersinia pestis* strains Antiqua and Nepal516: evidence of gene reduction in an emerging pathogen. *J Bacteriol* 188: 4453-4463.
- CHAPMAN MR, ROBINSON LS, PINKNER JS, ROTH R, HEUSER J, HAMMAR M, NORMARK S AND HULTGREN SJ. 2002. Role of *Escherichia coli* curli operons in directing amyloid fiber formation. *Science* 295: 851-855.
- CHASSIN C, TOURNEUR E, BENS M AND VANDEWALLE A. 2011. A role for collecting duct epithelial cells in renal antibacterial defences. *Cellular microbiology* 13: 1107-1113.
- CHEN SL ET AL. 2006. Identification of genes subject to positive selection in uropathogenic strains of *Escherichia coli*: a comparative genomics approach. *Proc Natl Acad Sci U S A* 103: 5977-5982.
- CHRIS DURAN NA, DAVID EDWARDS AND JACQUELINE BATLEY. 2009. Molecular Genetic Markers: Discovery, Applications, Data Storage and Visualisation. *Current Bioinformatics* 4: 16-27.
- CICCARELLI FD, DOERKS T, VON MERING C, CREEVEY CJ, SNEL B AND BORK P. 2006. Toward automatic reconstruction of a highly resolved tree of life. *Science* 311: 1283-1287.
- COLON-GONZALEZ M, MENDEZ-ORTIZ MM AND MEMBRILLO-HERNANDEZ J. 2004. Anaerobic growth does not support biofilm formation in *Escherichia coli* K-12. *Research in microbiology* 155: 514-521.
- COMAS I, HOMOLKA S, NIEMANN S AND GAGNEUX S. 2009. Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosis* highlights the limitations of current methodologies. *PLoS One* 4: e7815.
- CONNELL I, AGACE W, KLEMM P, SCHEMBRI M, MARILD S AND SVANBORG C. 1996. Type 1 fimbrial expression enhances *Escherichia coli* virulence for the urinary tract. *Proc Natl Acad Sci U S A* 93: 9827-9832.
- COOKSON AL, COOLEY WA AND WOODWARD MJ. 2002. The role of type 1 and curli fimbriae of Shiga toxin-producing *Escherichia coli* in adherence to abiotic surfaces. *Int J Med Microbiol* 292: 195-205.
- COOTE JG. 1995. The RTX toxins of Gram-negative bacterial pathogens: modulators of the host immune system. *Reviews in Medical Microbiology* 7(1).
- COQUE TM, NOVAIS A, CARATTOLI A, POIREL L, PITOUT J, PEIXE L, BAQUERO F, CANTON R AND NORDMANN P. 2008. Dissemination of clonally related *Escherichia coli* strains expressing extended-spectrum beta-lactamase CTX-M-15. *Emerging infectious diseases* 14: 195-200.
- CORANDER J, WALDMANN P AND SILLANPAA MJ. 2003. Bayesian analysis of genetic differentiation between populations. *Genetics* 163: 367-374.
- CORTES P ET AL. 2010. Isolation and characterization of potentially pathogenic antimicrobial-resistant *Escherichia coli* strains from chicken and pig farms in Spain. *Appl Environ Microbiol* 76: 2799-2805.
- COSTERTON JW. 1999. Introduction to biofilm. *Int J Antimicrob Agents* 11: 217-221; discussion 237-219.
- CROSS AS. 1990. The biologic significance of bacterial encapsulation. *Current topics in microbiology and immunology* 150: 87-95.

- CROXEN MA AND FINLAY BB. 2010. Molecular mechanisms of *Escherichia coli* pathogenicity. *Nat Rev Microbiol* 8: 26-38.
- CZIROK E, MILCH H, CSISZAR K AND CSIK M. 1986. Virulence factors of *Escherichia coli*. III. Correlation with *Escherichia coli* pathogenicity of haemolysin production, haemagglutinating capacity, antigens K1, K5, and colicinogenicity. *Acta microbiologica Hungarica* 33: 69-83.
- DARFEUILLE-MICHAUD A, BOUDEAU J, BULOIS P, NEUT C, GLASSER AL, BARNICH N, BRINGER MA, SWIDSINSKI A, BEAUGERIE L AND COLOMBEL JF. 2004. High prevalence of adherent-invasive *Escherichia coli* associated with ileal mucosa in Crohn's disease. *Gastroenterology* 127: 412-421.
- DAUBIN V, GOUY M AND PERRIERE G. 2002. A phylogenomic approach to bacterial phylogeny: evidence of a core of genes sharing a common history. *Genome research* 12: 1080-1090.
- DAVEY JW, HOHENLOHE PA, ETTER PD, BOONE JQ, CATCHEN JM AND BLAXTER ML. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature reviews Genetics* 12: 499-510.
- DAVEY ME AND O'TOOLE G A. 2000. Microbial biofilms: from ecology to molecular genetics. *Microbiology and molecular biology reviews* : MMBR 64: 847-867.
- DEMIR M AND KALELI I. 2004. Production by *Escherichia coli* isolates of siderophore and other virulence factors and their pathogenic role in a cutaneous infection model. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* 10: 1011-1014.
- DEVULDER G, PEROUSE DE MONTCLOS M AND FLANDROIS JP. 2005. A multigene approach to phylogenetic analysis using the genus *Mycobacterium* as a model. *International journal of systematic and evolutionary microbiology* 55: 293-302.
- DHO-MOULIN M AND FAIRBROTHER JM. 1999. Avian pathogenic *Escherichia coli* (APEC). *Veterinary research* 30: 299-316.
- DOBRINDT U, AGERER F, MICHAELIS K, JANKA A, BUCHRIESER C, SAMUELSON M, SVANBORG C, GOTTSCHALK G, KARCH H AND HACKER J. 2003. Analysis of genome plasticity in pathogenic and commensal *Escherichia coli* isolates by use of DNA arrays. *J Bacteriol* 185: 1831-1840.
- DOBRINDT U, BLUM-OEHLER G, HARTSCH T, GOTTSCHALK G, RON EZ, FUNFSTUCK R AND HACKER J. 2001. S-Fimbria-encoding determinant *sfa(I)* is located on pathogenicity island III(536) of uropathogenic *Escherichia coli* strain 536. *Infect Immun* 69: 4248-4256.
- DOBRINDT U AND HACKER J. 2008. Targeting virulence traits: potential strategies to combat extraintestinal pathogenic *E. coli* infections. *Current opinion in microbiology* 11: 409-413.
- DOGAN B, KLAESSIG S, RISHNIW M, ALMEIDA RA, OLIVER SP, SIMPSON K AND SCHUKKEN YH. 2006. Adherent and invasive *Escherichia coli* are associated with persistent bovine mastitis. *Vet Microbiol* 116: 270-282.
- DONNENBERG MS AND NATARO JP. 1995. Methods for studying adhesion of diarrheagenic *Escherichia coli*. *Methods Enzymol* 253: 324-336.
- DOZOIS CM, DHO-MOULIN M, BREE A, FAIRBROTHER JM, DESAUTELS C AND CURTISS R, 3RD. 2000. Relationship between the Tsh autotransporter and pathogenicity of avian *Escherichia coli* and localization and analysis of the Tsh genetic region. *Infect Immun* 68: 4145-4154.

- DREUX N ET AL. 2013. Point mutations in FimH adhesin of Crohn's disease-associated adherent-invasive *Escherichia coli* enhance intestinal inflammatory response. *PLoS pathogens* 9: e1003141.
- DZIVA F AND STEVENS MP. 2008. Colibacillosis in poultry: unravelling the molecular basis of virulence of avian pathogenic *Escherichia coli* in their natural hosts. *Avian pathology : journal of the WVPA* 37: 355-366.
- EDELMAN S, LESKELA S, RON E, APAJALAHTI J AND KORHONEN TK. 2003. In vitro adhesion of an avian pathogenic *Escherichia coli* O78 strain to surfaces of the chicken intestinal tract and to ileal mucus. *Vet Microbiol* 91: 41-56.
- EDEN CS AND HANSSON HA. 1978. *Escherichia coli* pili as possible mediators of attachment to human urinary tract epithelial cells. *Infect Immun* 21: 229-237.
- EDGAR RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5: 113.
- EISENSTEIN BI AND JONES GW. 1988. The spectrum of infections and pathogenic mechanisms of *Escherichia coli*. *Advances in internal medicine* 33: 231-252.
- ELSINGHORST EA. 1994. Measurement of invasion by gentamicin resistance. *Methods Enzymol* 236: 405-420.
- ETO DS, JONES TA, SUNDSBAK JL AND MULVEY MA. 2007. Integrin-mediated host cell invasion by type 1-piliated uropathogenic *Escherichia coli*. *PLoS pathogens* 3: e100.
- EWERS C, ANTAO EM, DIEHL I, PHILIPP HC AND WIELER LH. 2009. Intestine and environment of the chicken as reservoirs for extraintestinal pathogenic *Escherichia coli* strains with zoonotic potential. *Appl Environ Microbiol* 75: 184-192.
- EWERS C, GROBBEL M, BETHE A, WIELER LH AND GUENTHER S. 2011. Extended-spectrum beta-lactamases-producing gram-negative bacteria in companion animals: action is clearly warranted! *Berl Munch Tierarztl Wochenschr* 124: 94-101.
- EWERS C, JANSSEN T, KIESSLING S, PHILIPP HC AND WIELER LH. 2004. Molecular epidemiology of avian pathogenic *Escherichia coli* (APEC) isolated from colisepticemia in poultry. *Vet Microbiol* 104: 91-101.
- EWERS C, JANSSEN T, KIESSLING S, PHILIPP HC AND WIELER LH. 2005. Rapid detection of virulence-associated genes in avian pathogenic *Escherichia coli* by multiplex polymerase chain reaction. *Avian diseases* 49: 269-273.
- EWERS C ET AL. 2007. Avian pathogenic, uropathogenic, and newborn meningitis-causing *Escherichia coli*: how closely related are they? *Int J Med Microbiol* 297: 163-176.
- FALBO V, FAMIGLIETTI M AND CAPRIOLI A. 1992. Gene block encoding production of cytotoxic necrotizing factor 1 and hemolysin in *Escherichia coli* isolates from extraintestinal infections. *Infect Immun* 60: 2182-2187.
- FAM N, LEFLON-GUIBOUT V, FOUAD S, ABOUL-FADL L, MARCON E, DESOUKY D, EL-DEFRAWY I, ABOU-AITTA A, KLENA J AND NICOLAS-CHANOINE MH. 2011. CTX-M-15-producing *Escherichia coli* clinical isolates in Cairo (Egypt), including isolates of clonal complex ST10 and clones ST131, ST73, and ST405 in both community and hospital settings. *Microbial drug resistance* 17: 67-73.
- FEIL EJ. 2004. Small change: keeping pace with microevolution. *Nat Rev Microbiol* 2: 483-495.
- FELMLEE T, PELLETT S, LEE EY AND WELCH RA. 1985. *Escherichia coli* hemolysin is released extracellularly without cleavage of a signal peptide. *J Bacteriol* 163: 88-93.
- FENG L ET AL. 2008. A recalibrated molecular clock and independent origins for the cholera pandemic clones. *PLoS One* 3: e4053.

- FERNANDEZ-BEROS ME, KISSEL V, LIOR H AND CABELLO FC. 1990. Virulence-related genes in ColV plasmids of *Escherichia coli* isolated from human blood and intestines. *Journal of clinical microbiology* 28: 742-746.
- FETHERSTON JD AND PERRY RD. 1994. The pigmentation locus of *Yersinia pestis* KIM6+ is flanked by an insertion sequence and includes the structural genes for pesticin sensitivity and HMWP2. *Mol Microbiol* 13: 697-708.
- FINLAY BB AND FALKOW S. 1997. Common themes in microbial pathogenicity revisited. *Microbiology and molecular biology reviews* : MMBR 61: 136-169.
- FISCHBACH MA ET AL. 2006. The pathogen-associated iroA gene cluster mediates bacterial evasion of lipocalin 2. *Proc Natl Acad Sci U S A* 103: 16502-16507.
- FISCHER H, YAMAMOTO M, AKIRA S, BEUTLER B AND SVANBORG C. 2006. Mechanism of pathogen-specific TLR4 activation in the mucosa: fimbriae, recognition receptors and adaptor protein selection. *European journal of immunology* 36: 267-277.
- FORGETTA V ET AL. 2011. Fourteen-genome comparison identifies DNA markers for severe-disease-associated strains of *Clostridium difficile*. *Journal of clinical microbiology* 49: 2230-2238.
- FOXMAN B. 2002. Epidemiology of urinary tract infections: incidence, morbidity, and economic costs. *The American journal of medicine* 113 Suppl 1A: 5S-13S.
- FOXMAN B, ZHANG L, TALLMAN P, PALIN K, RODE C, BLOCH C, GILLESPIE B AND MARRS CF. 1995. Virulence characteristics of *Escherichia coli* causing first urinary tract infection predict risk of second infection. *The Journal of infectious diseases* 172: 1536-1541.
- GODALY G ET AL. 2000. Innate defences and resistance to gram negative mucosal infection. *Adv Exp Med Biol* 485: 9-24.
- GOEBEL W AND HEDGPETH J. 1982. Cloning and functional characterization of the plasmid-encoded hemolysin determinant of *Escherichia coli*. *J Bacteriol* 151: 1290-1298.
- GOETZ DH, HOLMES MA, BORREGAARD N, BLUHM ME, RAYMOND KN AND STRONG RK. 2002. The neutrophil lipocalin NGAL is a bacteriostatic agent that interferes with siderophore-mediated iron acquisition. *Molecular cell* 10: 1033-1043.
- GOLDMAN RC, JOINER K AND LEIVE L. 1984. Serum-resistant mutants of *Escherichia coli* O111 contain increased lipopolysaccharide, lack an O antigen-containing capsule, and cover more of their lipid A core with O antigen. *J Bacteriol* 159: 877-882.
- GOPHNA U, BARLEV M, SEIFFERS R, OELSCHLAGER TA, HACKER J AND RON EZ. 2001. Curli fibers mediate internalization of *Escherichia coli* by eukaryotic cells. *Infect Immun* 69: 2659-2665.
- GRODBERG J, LUNDRIGAN MD, TOLEDO DL, MANGEL WF AND DUNN JJ. 1988. Complete nucleotide sequence and deduced amino acid sequence of the *ompT* gene of *Escherichia coli* K-12. *Nucleic Acids Res* 16: 1209.
- GROISMAN EA AND OCHMAN H. 1996. Pathogenicity islands: bacterial evolution in quantum leaps. *Cell* 87: 791-794.
- GUALDI L, TAGLIABUE L, BERTAGNOLI S, IERANO T, DE CASTRO C AND LANDINI P. 2008. Cellulose modulates biofilm formation by counteracting curli-mediated colonization of solid surfaces in *Escherichia coli*. *Microbiology* 154: 2017-2024.
- GUENTHER S, BETHE A, FRUTH A, SEMMLER T, ULRICH RG, WIELER LH AND EWERS C. 2012. Frequent combination of antimicrobial multiresistance and extraintestinal pathogenicity in *Escherichia coli* isolates from urban rats (*Rattus norvegicus*) in Berlin, Germany. *PLoS One* 7: e50331.

- GUINDON S, DUFAYARD JF, LEFORT V, ANISIMOVA M, HORDIJK W AND GASCUEL O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 59: 307-321.
- GUNTHER NWT, SNYDER JA, LOCKATELL V, BLOMFIELD I, JOHNSON DE AND MOBLEY HL. 2002. Assessment of virulence of uropathogenic *Escherichia coli* type 1 fimbrial mutants in which the invertible element is phase-locked on or off. *Infect Immun* 70: 3344-3354.
- GUPTA PK RJ, PRASAD M. 2001. Single nucleotide polymorphisms: A new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. *Curr Sci* 80: 524-535.
- GUTTMAN DS DD. 1994. Clonal divergence in *Escherichia coli* as a result of recombination, not mutation. *Science* 266(5189): 1380-1383.
- GUYER DM, HENDERSON IR, NATARO JP AND MOBLEY HL. 2000. Identification of sat, an autotransporter toxin produced by uropathogenic *Escherichia coli*. *Mol Microbiol* 38: 53-66.
- GUYER DM, KAO JS AND MOBLEY HL. 1998. Genomic analysis of a pathogenicity island in uropathogenic *Escherichia coli* CFT073: distribution of homologous sequences among isolates from patients with pyelonephritis, cystitis, and Catheter-associated bacteriuria and from fecal samples. *Infect Immun* 66: 4411-4417.
- GUYER DM, RADULOVIC S, JONES FE AND MOBLEY HL. 2002. Sat, the secreted autotransporter toxin of uropathogenic *Escherichia coli*, is a vacuolating cytotoxin for bladder and kidney epithelial cells. *Infect Immun* 70: 4539-4546.
- GYLES. 1994. Diseases due to *Escherichia coli* in poultry: *Escherichia coli* in domestic animals and man. CAB International, Wallingford, United Kingdom: 237-259.
- HABILITATION-EWERS. 2012. Phylogenetische Untersuchungen extraintestinal pathogener *Escherichia coli* (ExPEC) von Tier und Mensch als Grundlage zur Bewertung der Virulenz, klinischen Bedeutung und des Zoonosepotentials aviärer pathogener *E. coli* (APEC). unpublished
- HACKER J AND CARNIEL E. 2001. Ecological fitness, genomic islands and bacterial pathogenicity. A Darwinian view of the evolution of microbes. *EMBO Rep* 2: 376-381.
- HAGBERG L, JODAL U, KORHONEN TK, LIDIN-JANSON G, LINDBERG U AND SVANBORG EDEN C. 1981. Adhesion, hemagglutination, and virulence of *Escherichia coli* causing urinary tract infections. *Infect Immun* 31: 564-570.
- HAHN E, WILD P, HERMANN S, SEBBEL P, GLOCKSHUBER R, HANER M, TASCHNER N, BURKHARD P, AEBI U AND MULLER SA. 2002. Exploring the 3D molecular architecture of *Escherichia coli* type 1 pili. *J Mol Biol* 323: 845-857.
- HANNA A, BERG M, STOUT V AND RAZATOS A. 2003. Role of capsular colanic acid in adhesion of uropathogenic *Escherichia coli*. *Appl Environ Microbiol* 69: 4474-4481.
- HANTKE K, NICHOLSON G, RABSCH W AND WINKELMANN G. 2003. Salmochelins, siderophores of *Salmonella enterica* and uropathogenic *Escherichia coli* strains, are recognized by the outer membrane receptor IroN. *Proc Natl Acad Sci U S A* 100: 3677-3682.
- HEDLUND M, WACHTLER C, JOHANSSON E, HANG L, SOMERVILLE JE, DARVEAU RP AND SVANBORG C. 1999. P fimbriae-dependent, lipopolysaccharide-independent activation of epithelial cytokine responses. *Mol Microbiol* 33: 693-703.
- HENDERSON IR, CZECZULIN J, ESLAVA C, NORIEGA F AND NATARO JP. 1999. Characterization of *pic*, a secreted protease of *Shigella flexneri* and enteroaggregative *Escherichia coli*. *Infect Immun* 67: 5587-5596.

- HOFFMANN C AND SCHMIDT G. 2004. CNF and DNT. Reviews of physiology, biochemistry and pharmacology 152: 49-63.
- HOFMAN P, LE NEGRATE G, MOGRABI B, HOFMAN V, BREST P, ALLIANA-SCHMID A, FLATAU G, BOQUET P AND ROSSI B. 2000. *Escherichia coli* cytotoxic necrotizing factor-1 (CNF-1) increases the adherence to epithelia and the oxidative burst of human polymorphonuclear leukocytes but decreases bacteria phagocytosis. Journal of leukocyte biology 68: 522-528.
- HOLDEN NJ AND GALLY DL. 2004. Switches, cross-talk and memory in *Escherichia coli* adherence. J Med Microbiol 53: 585-593.
- HOLMES EC, URWIN R AND MAIDEN MC. 1999. The influence of recombination on the population structure and evolution of the human pathogen *Neisseria meningitidis*. Mol Biol Evol 16: 741-749.
- HOMMAIS F, GOURIOU S, AMORIN C, BUI H, RAHIMY MC, PICARD B AND DENAMUR E. 2003. The FimH A27V mutation is pathoadaptive for urovirulence in *Escherichia coli* B2 phylogenetic group isolates. Infect Immun 71: 3619-3622.
- HULL RA, GILL RE, HSU P, MINSHEW BH AND FALKOW S. 1981. Construction and expression of recombinant plasmids encoding type 1 or D-mannose-resistant pili from a urinary tract infection *Escherichia coli* isolate. Infect Immun 33: 933-938.
- HUSSAIN A, EWERS C, NANDANWAR N, GUENTHER S, JADHAV S, WIELER LH AND AHMED N. 2012. Multiresistant uropathogenic *Escherichia coli* from a region in India where urinary tract infections are endemic: genotypic and phenotypic characteristics of sequence type 131 isolates of the CTX-M-15 extended-spectrum-beta-lactamase-producing lineage. Antimicrobial agents and chemotherapy 56: 6358-6365.
- IDESES D, GOPHNA U, PAITAN Y, CHAUDHURI RR, PALLEN MJ AND RON EZ. 2005. A degenerate type III secretion system from septicemic *Escherichia coli* contributes to pathogenesis. J Bacteriol 187: 8164-8171.
- JACKSON DW, SIMECKA JW AND ROMEO T. 2002. Catabolite repression of *Escherichia coli* biofilm formation. J Bacteriol 184: 3406-3410.
- JANN K AND JANN B. 1983. The K antigens of *Escherichia coli*. Progress in allergy 33: 53-79.
- JANN K AND JANN B. 1987. Polysaccharide antigens of *Escherichia coli*. Rev Infect Dis 9 Suppl 5: S517-526.
- JANßEN T, SCHWARZ C, PREIKSCHAT P, VOSS M, PHILIPP HC AND WIELER LH. 2001. Virulence-associated genes in avian pathogenic *Escherichia coli* (APEC) isolated from internal organs of poultry having died from colibacillosis. Int J Med Microbiol 291: 371-378.
- JANTAUSCH BA AND HULL SI. 1996. Restriction fragment length polymorphism of PCR amplified papE gene products is correlated with complete serotype among uropathogenic *Escherichia coli* isolates. Microb Pathog 20: 351-360.
- JOHNSON JR. 1991. Virulence factors in *Escherichia coli* urinary tract infection. Clin Microbiol Rev 4: 80-128.
- JOHNSON JR, CLABOTS C AND KUSKOWSKI MA. 2008a. Multiple-host sharing, long-term persistence, and virulence of *Escherichia coli* clones from human and animal household members. Journal of clinical microbiology 46: 4078-4082.
- JOHNSON JR, CLABOTS C AND ROSEN H. 2006. Effect of inactivation of the global oxidative stress regulator oxyR on the colonization ability of *Escherichia coli* O1:K1:H7 in a mouse model of ascending urinary tract infection. Infect Immun 74: 461-468.

- JOHNSON JR, DELAVARI P, STELL AL, WHITTAM TS, CARLINO U AND RUSSO TA. 2001a. Molecular comparison of extraintestinal *Escherichia coli* isolates of the same electrophoretic lineages from humans and domestic animals. *The Journal of infectious diseases* 183: 154-159.
- JOHNSON JR, JELACIC S, SCHOENING LM, CLABOTS C, SHAIKH N, MOBLEY HL AND TARR PI. 2005a. The IrgA homologue adhesin Iha is an *Escherichia coli* virulence factor in murine urinary tract infection. *Infect Immun* 73: 965-971.
- JOHNSON JR, JOHNSTON B, CLABOTS CR, KUSKOWSKI MA, ROBERTS E AND DEBROY C. 2008b. Virulence genotypes and phylogenetic background of *Escherichia coli* serogroup O6 isolates from humans, dogs, and cats. *Journal of clinical microbiology* 46: 417-422.
- JOHNSON JR, KUSKOWSKI MA, O'BRYAN TT AND MASLOW JN. 2002a. Epidemiological correlates of virulence genotype and phylogenetic background among *Escherichia coli* blood isolates from adults with diverse-source bacteremia. *The Journal of infectious diseases* 185: 1439-1447.
- JOHNSON JR, MILLER S, JOHNSTON B, CLABOTS C AND DEBROY C. 2009. Sharing of *Escherichia coli* sequence type ST131 and other multidrug-resistant and Urovirulent *E. coli* strains among dogs and cats within a household. *Journal of clinical microbiology* 47: 3721-3725.
- JOHNSON JR, MURRAY AC, GAJEWSKI A, SULLIVAN M, SNIPPES P, KUSKOWSKI MA AND SMITH KE. 2003. Isolation and molecular characterization of nalidixic acid-resistant extraintestinal pathogenic *Escherichia coli* from retail chicken products. *Antimicrobial agents and chemotherapy* 47: 2161-2168.
- JOHNSON JR, O'BRYAN TT, LOW DA, LING G, DELAVARI P, FASCHING C, RUSSO TA, CARLINO U AND STELL AL. 2000. Evidence of commonality between canine and human extraintestinal pathogenic *Escherichia coli* strains that express *papG* allele III. *Infect Immun* 68: 3327-3336.
- JOHNSON JR, OSWALD E, O'BRYAN TT, KUSKOWSKI MA AND SPANJAARD L. 2002b. Phylogenetic distribution of virulence-associated genes among *Escherichia coli* isolates associated with neonatal bacterial meningitis in the Netherlands. *The Journal of infectious diseases* 185: 774-784.
- JOHNSON JR AND RUSSO TA. 2002. Extraintestinal pathogenic *Escherichia coli*: "the other bad *E. coli*". *The Journal of laboratory and clinical medicine* 139: 155-162.
- JOHNSON JR AND RUSSO TA. 2005. Molecular epidemiology of extraintestinal pathogenic (uropathogenic) *Escherichia coli*. *Int J Med Microbiol* 295: 383-404.
- JOHNSON JR, SCHEUTZ F, ULLERYD P, KUSKOWSKI MA, O'BRYAN TT AND SANDBERG T. 2005b. Phylogenetic and pathotypic comparison of concurrent urine and rectal *Escherichia coli* isolates from men with febrile urinary tract infection. *Journal of clinical microbiology* 43: 3895-3900.
- JOHNSON JR AND STELL AL. 2000. Extended virulence genotypes of *Escherichia coli* strains from patients with urosepsis in relation to phylogeny and host compromise. *The Journal of infectious diseases* 181: 261-272.
- JOHNSON JR, STELL AL, DELAVARI P, MURRAY AC, KUSKOWSKI M AND GAASTRA W. 2001b. Phylogenetic and pathotypic similarities between *Escherichia coli* isolates from urinary tract infections in dogs and extraintestinal infections in humans. *The Journal of infectious diseases* 183: 897-906.
- JOHNSON TJ, KARIYAWASAM S, WANNEMUEHLER Y, MANGIAMELE P, JOHNSON SJ, DOETKOTT C, SKYBERG JA, LYNNE AM, JOHNSON JR AND NOLAN LK.

2007. The genome sequence of avian pathogenic *Escherichia coli* strain O1:K1:H7 shares strong similarities with human extraintestinal pathogenic *E. coli* genomes. *J Bacteriol* 189: 3228-3236.
- JOHNSON TJ, WANNEMUEHLER Y, DOETKOTT C, JOHNSON SJ, ROSENBERGER SC AND NOLAN LK. 2008c. Identification of minimal predictors of avian pathogenic *Escherichia coli* virulence for use as a rapid diagnostic tool. *Journal of clinical microbiology* 46: 3987-3996.
- JOHNSON TJ, WANNEMUEHLER Y, JOHNSON SJ, STELL AL, DOETKOTT C, JOHNSON JR, KIM KS, SPANJAARD L AND NOLAN LK. 2008d. Comparison of extraintestinal pathogenic *Escherichia coli* strains from human and avian sources reveals a mixed subset representing potential zoonotic pathogens. *Appl Environ Microbiol* 74: 7043-7050.
- JONES CH, JACOB-DUBUISSON F, DODSON K, KUEHN M, SLONIM L, STRIKER R AND HULTGREN SJ. 1992. Adhesin presentation in bacteria requires molecular chaperones and ushers. *Infect Immun* 60: 4445-4451.
- JUNEMANN S ET AL. 2013. Updating benchtop sequencing performance comparison. *Nature biotechnology* 31: 294-296.
- JUSTICE SS, LAUER SR, HULTGREN SJ AND HUNSTAD DA. 2006. Maturation of intracellular *Escherichia coli* communities requires SurA. *Infect Immun* 74: 4793-4800.
- KAAS RS, FRIIS C, USSERY DW AND AARESTRUP FM. 2012. Estimating variation within the genes and inferring the phylogeny of 186 sequenced diverse *Escherichia coli* genomes. *BMC genomics* 13: 577.
- KANAMARU S, KURAZONO H, ISHITOYA S, TERAJ A, HABUCHI T, NAKANO M, OGAWA O AND YAMAMOTO S. 2003. Distribution and genetic association of putative uropathogenic virulence factors *iroN*, *iha*, *kpsMT*, *ompT* and *usp* in *Escherichia coli* isolated from urinary tract infections in Japan. *The Journal of urology* 170: 2490-2493.
- KAPER JB, NATARO JP AND MOBLEY HL. 2004. Pathogenic *Escherichia coli*. *Nat Rev Microbiol* 2: 123-140.
- KARIYAWASAM S, JOHNSON TJ, DEBROY C AND NOLAN LK. 2006. Occurrence of pathogenicity island I(APEC-O1) genes among *Escherichia coli* implicated in avian colibacillosis. *Avian diseases* 50: 405-410.
- KHAN AS, KNIEP B, OELSCHLAEGER TA, VAN DIE I, KORHONEN T AND HACKER J. 2000. Receptor structure for F1C fimbriae of uropathogenic *Escherichia coli*. *Infect Immun* 68: 3541-3547.
- KIM KS. 2001. *Escherichia coli* translocation at the blood-brain barrier. *Infect Immun* 69: 5217-5222.
- KIM KS. 2003. Pathogenesis of bacterial meningitis: from bacteraemia to neuronal injury. *Nature reviews Neuroscience* 4: 376-385.
- KLEMM P. 1986. Two regulatory fim genes, fimB and fimE, control the phase variation of type 1 fimbriae in *Escherichia coli*. *The EMBO journal* 5: 1389-1393.
- KLEMM P AND CHRISTIANSEN G. 1987. Three fim genes required for the regulation of length and mediation of adhesion of *Escherichia coli* type 1 fimbriae. *Molecular & general genetics* : MGG 208: 439-445.
- KLEMM P, CHRISTIANSEN G, KREFT B, MARRE R AND BERGMANS H. 1994. Reciprocal exchange of minor components of type 1 and F1C fimbriae results in hybrid organelles with changed receptor specificities. *J Bacteriol* 176: 2227-2234.

- KLEMM P, HANCOCK V AND SCHEMBRI MA. 2010. Fimbrial adhesins from extraintestinal *Escherichia coli*. Environmental microbiology reports 2: 628-640.
- KLEMM P, JORGENSEN BJ, KREFT B AND CHRISTIANSEN G. 1995. The export systems of type 1 and F1C fimbriae are interchangeable but work in parental pairs. J Bacteriol 177: 621-627.
- KLEMM P, JORGENSEN BJ, VAN DIE I, DE REE H AND BERGMANS H. 1985. The fim genes responsible for synthesis of type 1 fimbriae in *Escherichia coli*, cloning and genetic organization. Molecular & general genetics : MGG 199: 410-414.
- KLEMM P AND SCHEMBRI MA. 2000. Bacterial adhesins: function and structure. Int J Med Microbiol 290: 27-35.
- KLOSE M, STORIKO A, STIERHOF YD, HINDENNACH I, MUTSCHLER B AND HENNING U. 1993. Membrane assembly of the outer membrane protein OmpA of *Escherichia coli*. J Biol Chem 268: 25664-25670.
- KOCZURA R AND KAZNOWSKI A. 2003. Occurrence of the *Yersinia* high-pathogenicity island and iron uptake systems in clinical isolates of *Klebsiella pneumoniae*. Microb Pathog 35: 197-202.
- KOLCHINSKY AM, FUNKE RP AND GRESSHOFF PM. 1993. DAF-amplified fragments can be used as markers for DNA from pulse field gels. BioTechniques 14: 400-403.
- KONIECZNY A AND AUSUBEL FM. 1993. A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. The Plant journal : for cell and molecular biology 4: 403-410.
- KORHONEN TK, PARKKINEN J, HACKER J, FINNE J, PERE A, RHEN M AND HOLTHOFER H. 1986. Binding of *Escherichia coli* S fimbriae to human kidney epithelium. Infect Immun 54: 322-327.
- KORHONEN TK, VAISANEN-RHEN V, RHEN M, PERE A, PARKKINEN J AND FINNE J. 1984. *Escherichia coli* fimbriae recognizing sialyl galactosides. J Bacteriol 159: 762-766.
- KROGFELT KA, BERGMANS H AND KLEMM P. 1990. Direct evidence that the FimH protein is the mannose-specific adhesin of *Escherichia coli* type 1 fimbriae. Infect Immun 58: 1995-1998.
- KUEHN MJ, HEUSER J, NORMARK S AND HULTGREN SJ. 1992. P pili in uropathogenic *E. coli* are composite fibres with distinct fibrillar adhesive tips. Nature 356: 252-255.
- KUKKONEN M AND KORHONEN TK. 2004. The omptin family of enterobacterial surface proteases/adhesins: from housekeeping in *Escherichia coli* to systemic spread of *Yersinia pestis*. Int J Med Microbiol 294: 7-14.
- LABIGNE-ROUSSEL A, SCHMIDT MA, WALZ W AND FALKOW S. 1985. Genetic organization of the afimbrial adhesin operon and nucleotide sequence from a uropathogenic *Escherichia coli* gene encoding an afimbrial adhesin. J Bacteriol 162: 1285-1292.
- LABIGNE-ROUSSEL AF, LARK D, SCHOOLNIK G AND FALKOW S. 1984. Cloning and expression of an afimbrial adhesin (AFA-I) responsible for P blood group-independent, mannose-resistant hemagglutination from a pyelonephritic *Escherichia coli* strain. Infect Immun 46: 251-259.
- LANDRAUD L, PULCINI C, GOUNON P, FLATAU G, BOQUET P AND LEMICHEZ E. 2004. *E. coli* CNF1 toxin: a two-in-one system for host-cell invasion. Int J Med Microbiol 293: 513-518.
- LARKIN MA ET AL. 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947-2948.

- LASARO MA, SALINGER N, ZHANG J, WANG Y, ZHONG Z, GOULIAN M AND ZHU J. 2009. F1C fimbriae play an important role in biofilm formation and intestinal colonization by the *Escherichia coli* commensal strain Nissle 1917. *Appl Environ Microbiol* 75: 246-251.
- LAU SH, REDDY S, CHEESBROUGH J, BOLTON FJ, WILLSHAW G, CHEASTY T, FOX AJ AND UPTON M. 2008. Major uropathogenic *Escherichia coli* strain isolated in the northwest of England identified by multilocus sequence typing. *Journal of clinical microbiology* 46: 1076-1080.
- LEMICHEZ E, FLATAU G, BRUZZONE M, BOQUET P AND GAUTHIER M. 1997. Molecular localization of the *Escherichia coli* cytotoxic necrotizing factor CNF1 cell-binding and catalytic domains. *Mol Microbiol* 24: 1061-1070.
- LEYING H, SUERBAUM S, KROLL HP, STAHL D AND OPFERKUCH W. 1990. The capsular polysaccharide is a major determinant of serum resistance in K-1-positive blood culture isolates of *Escherichia coli*. *Infect Immun* 58: 222-227.
- LI G, CAI W, HUSSEIN A, WANNEMUEHLER YM, LOGUE CM AND NOLAN LK. 2012. Proteome response of an extraintestinal pathogenic *Escherichia coli* strain with zoonotic potential to human and chicken sera. *Journal of proteomics* 75: 4853-4862.
- LI G, LATURNUS C, EWERS C AND WIELER LH. 2005. Identification of genes required for avian *Escherichia coli* septicemia by signature-tagged mutagenesis. *Infect Immun* 73: 2818-2827.
- LI G ET AL. 2011. Transcriptome analysis of avian pathogenic *Escherichia coli* O1 in chicken serum reveals adaptive responses to systemic infection. *Infect Immun* 79: 1951-1960.
- MABBETT AN ET AL. 2009. Virulence properties of asymptomatic bacteriuria *Escherichia coli*. *Int J Med Microbiol* 299: 53-63.
- MAIDEN MC ET AL. 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci U S A* 95: 3140-3145.
- MALORNI W, QUARANTA MG, STRAFACE E, FALZANO L, FABBRI A, VIORA M AND FIORENTINI C. 2003. The Rac-activating toxin cytotoxic necrotizing factor 1 oversees NK cell-mediated activity by regulating the actin/microtubule interplay. *J Immunol* 171: 4195-4202.
- MARGARITA PN, JORGE BLANCO³, XAVIER ALDEGUER⁴, JESUS E BLANCO³ MB, CARMEN PONTE², FRANCISCO SORIANO², AND GARCIA-GIL*¹ AD-MALJ. 2009. Biofilm formation as a novel phenotypic feature of adherent-invasive *Escherichia coli* (AIEC). *BMC Microbiology* 9.
- MARRS CF, ZHANG L AND FOXMAN B. 2005. *Escherichia coli* mediated urinary tract infections: are there distinct uropathogenic *E. coli* (UPEC) pathotypes? *FEMS Microbiol Lett* 252: 183-190.
- MARRS CF ET AL. 2002. Variations in 10 putative uropathogen virulence genes among urinary, faecal and peri-urethral *Escherichia coli*. *J Med Microbiol* 51: 138-142.
- MARTINEZ-MEDINA M, GARCIA-GIL J, BARNICH N, WIELER LH AND EWERS C. 2011. Adherent-invasive *Escherichia coli* phenotype displayed by intestinal pathogenic *E. coli* strains from cats, dogs, and swine. *Appl Environ Microbiol* 77: 5813-5817.
- MARTINEZ JJ, MULVEY MA, SCHILLING JD, PINKNER JS AND HULTGREN SJ. 2000. Type 1 pilus-mediated bacterial invasion of bladder epithelial cells. *The EMBO journal* 19: 2803-2812.

- MATTER LB, BARBIERI NL, NORDHOFF M, EWERS C AND HORN F. 2011. Avian pathogenic *Escherichia coli* MT78 invades chicken fibroblasts. *Vet Microbiol* 148: 51-59.
- MAUCHLINE TH, KNOX R, MOHAN S, POWERS SJ, KERRY BR, DAVIES KG AND HIRSCH PR. 2011. Identification of new single nucleotide polymorphism-based markers for inter- and intraspecies discrimination of obligate bacterial parasites (*Pasteuria spp.*) of invertebrates. *Appl Environ Microbiol* 77: 6388-6394.
- MEDOF ME, WALTER EI, RUTGERS JL, KNOWLES DM AND NUSSENZWEIG V. 1987. Identification of the complement decay-accelerating factor (DAF) on epithelium and glandular cells and in body fluids. *The Journal of experimental medicine* 165: 848-864.
- MELLATA M, DHO-MOULIN M, DOZOIS CM, CURTISS R, 3RD, BROWN PK, ARNE P, BREE A, DESAUTELS C AND FAIRBROTHER JM. 2003. Role of virulence factors in resistance of avian pathogenic *Escherichia coli* to serum and in pathogenicity. *Infect Immun* 71: 536-540.
- MIETZNER TA AND MORSE SA. 1994. The role of iron-binding proteins in the survival of pathogenic bacteria. *Annual review of nutrition* 14: 471-493.
- MILLS M, MEYSICK KC AND O'BRIEN AD. 2000. Cytotoxic necrotizing factor type 1 of uropathogenic *Escherichia coli* kills cultured human uroepithelial 5637 cells by an apoptotic mechanism. *Infect Immun* 68: 5869-5880.
- MOCH T, HOSCHUTZKY H, HACKER J, KRONCKE KD AND JANN K. 1987. Isolation and characterization of the alpha-sialyl-beta-2,3-galactosyl-specific adhesin from fimbriated *Escherichia coli*. *Proc Natl Acad Sci U S A* 84: 3462-3466.
- MOKADY D, GOPHNA U AND RON EZ. 2005. Virulence factors of septicemic *Escherichia coli* strains. *Int J Med Microbiol* 295: 455-462.
- MORA A ET AL. 2009. Extraintestinal pathogenic *Escherichia coli* O1:K1:H7/NM from human and avian origin: detection of clonal groups B2 ST95 and D ST59 with different host distribution. *BMC Microbiol* 9: 132.
- MORIEL DG ET AL. 2010. Identification of protective and broadly conserved vaccine antigens from the genome of extraintestinal pathogenic *Escherichia coli*. *Proc Natl Acad Sci U S A* 107: 9072-9077.
- MOULIN-SCHOULEUR M RM, LAURENT S, BRÉE A, MIGNON-GRASTEAU S, GERMON P, RASSCHAERT D, SCHOULER C. 2007. Extraintestinal pathogenic *Escherichia coli* strains of avian and human origin: link between phylogenetic relationships and common virulence patterns. *J Clin Microbiol* 45(10): 3366-3376.
- MULLER CM, ABERG A, STRASEVICIENE J, EMODY L, UHLIN BE AND BALSALOBRE C. 2009. Type 1 fimbriae, a colonization factor of uropathogenic *Escherichia coli*, are controlled by the metabolic sensor CRP-cAMP. *PLoS pathogens* 5: e1000303.
- MULLER D, GREUNE L, HEUSIPP G, KARCH H, FRUTH A, TSCHAPE H AND SCHMIDT MA. 2007. Identification of unconventional intestinal pathogenic *Escherichia coli* isolates expressing intermediate virulence factor profiles by using a novel single-step multiplex PCR. *Appl Environ Microbiol* 73: 3380-3390.
- MULVEY MA, SCHILLING JD AND HULTGREN SJ. 2001. Establishment of a persistent *Escherichia coli* reservoir during the acute phase of a bladder infection. *Infect Immun* 69: 4572-4579.
- MUNRO P, FLATAU G, DOYE A, BOYER L, OREGIONI O, MEGE JL, LANDRAUD L AND LEMICHEZ E. 2004. Activation and proteasomal degradation of rho GTPases by

- cytotoxic necrotizing factor-1 elicit a controlled inflammatory response. *J Biol Chem* 279: 35849-35857.
- MURASE K ET AL. 2012. Haemolysin E- and enterohaemolysin-derived haemolytic activity of O55/O157 strains and other *Escherichia coli* lineages. *Microbiology* 158: 746-758.
- NAGY G AA, KNAPP O, MAIER E, DOBRINDT U, BLUM-OEHLER G, BENZ R, EMODY L, HACKER J. 2006. Both alpha-haemolysin determinants contribute to full virulence of uropathogenic *Escherichia coli* strain 536. *Microbes Infect* 8: 2006-2012.
- NATARO JP AND KAPER JB. 1998. Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev* 11: 142-201.
- NEGRE VL, BONACORSI S, SCHUBERT S, BIDET P, NASSIF X AND BINGEN E. 2004. The siderophore receptor IroN, but not the high-pathogenicity island or the hemin receptor ChuA, contributes to the bacteremic step of *Escherichia coli* neonatal meningitis. *Infect Immun* 72: 1216-1220.
- NIELSEN R, PAUL JS, ALBRECHTSEN A AND SONG YS. 2011. Genotype and SNP calling from next-generation sequencing data. *Nature reviews Genetics* 12: 443-451.
- NILIUS AM AND SAVAGE DC. 1984. Serum resistance encoded by colicin V plasmids in *Escherichia coli* and its relationship to the plasmid transfer system. *Infect Immun* 43: 947-953.
- NOEGEL A, RDEST U AND GOEBEL W. 1981. Determination of the functions of hemolytic plasmid pHly152 of *Escherichia coli*. *J Bacteriol* 145: 233-247.
- NOLAN LK, HORNE SM, GIDDINGS CW, FOLEY SL, JOHNSON TJ, LYNNE AM AND SKYBERG J. 2003. Resistance to serum complement, iss, and virulence of avian *Escherichia coli*. *Veterinary research communications* 27: 101-110.
- NOUGAYREDE JP, TAIEB F, DE RYCKE J AND OSWALD E. 2005. Cyclomodulins: bacterial effectors that modulate the eukaryotic cell cycle. *Trends Microbiol* 13: 103-110.
- NOWICKI B, SVANBORG-EDEN C, HULL R AND HULL S. 1989. Molecular analysis and epidemiology of the Dr hemagglutinin of uropathogenic *Escherichia coli*. *Infect Immun* 57: 446-451.
- OELSCHLAEGER TA, DOBRINDT U AND HACKER J. 2002. Virulence factors of uropathogens. *Current opinion in urology* 12: 33-38.
- OLSEN A, ARNQVIST A, HAMMAR M, SUKUPOLVI S AND NORMARK S. 1993. The RpoS sigma factor relieves H-NS-mediated transcriptional repression of *csgA*, the subunit gene of fibronectin-binding curli in *Escherichia coli*. *Mol Microbiol* 7: 523-536.
- OLSEN A, JONSSON A AND NORMARK S. 1989. Fibronectin binding mediated by a novel class of surface organelles on *Escherichia coli*. *Nature* 338: 652-655.
- OLSEN A, WICK MJ, MORGELIN M AND BJORCK L. 1998. Curli, fibrous surface proteins of *Escherichia coli*, interact with major histocompatibility complex class I molecules. *Infect Immun* 66: 944-949.
- OLSEN GJ, MATSUDA H, HAGSTROM R AND OVERBEEK R. 1994. fastDNAmL: a tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. *Comput Appl Biosci* 10: 41-48.
- ONG CL, ULETT GC, MABBETT AN, BEATSON SA, WEBB RI, MONAGHAN W, NIMMO GR, LOOKE DF, MCEWAN AG AND SCHEMBRI MA. 2008. Identification of type 3 fimbriae in uropathogenic *Escherichia coli* reveals a role in biofilm formation. *J Bacteriol* 190: 1054-1063.

- OROPEZA-WEKERLE RL, MULLER E, KERN P, MEYERMANN R AND GOEBEL W. 1989. Synthesis, inactivation, and localization of extracellular and intracellular *Escherichia coli* hemolysins. *J Bacteriol* 171: 2783-2788.
- ORSKOV I, ORSKOV F, BIRCH-ANDERSEN A, KANAMORI M AND SVANBORG-EDEN C. 1982. O, K, H and fimbrial antigens in *Escherichia coli* serotypes associated with pyelonephritis and cystitis. *Scandinavian journal of infectious diseases Supplementum* 33: 18-25.
- OSWALD E, SUGAI M, LABIGNE A, WU HC, FIORENTINI C, BOQUET P AND O'BRIEN AD. 1994. Cytotoxic necrotizing factor type 2 produced by virulent *Escherichia coli* modifies the small GTP-binding proteins Rho involved in assembly of actin stress fibers. *Proc Natl Acad Sci U S A* 91: 3814-3818.
- OTTO BR, VAN DOOREN SJ, NUIJENS JH, LUIRINK J AND OUDEGA B. 1998. Characterization of a hemoglobin protease secreted by the pathogenic *Escherichia coli* strain EB1. *The Journal of experimental medicine* 188: 1091-1103.
- OXHAMRE C, RICHTER-DAHLFORS A, ZHDANOV VP AND KASEMO B. 2005. A minimal generic model of bacteria-induced intracellular Ca²⁺ oscillations in epithelial cells. *Biophys J* 88: 2976-2981.
- PAGE RD. 1996. TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* 12: 357-358.
- PARHAM NJ, SRINIVASAN U, DESVAUX M, FOXMAN B, MARRS CF AND HENDERSON IR. 2004. PicU, a second serine protease autotransporter of uropathogenic *Escherichia coli*. *FEMS Microbiol Lett* 230: 73-83.
- PHAM TQ, GOLUSZKO P, POPOV V, NOWICKI S AND NOWICKI BJ. 1997. Molecular cloning and characterization of Dr-II, a nonfimbrial adhesin-I-like adhesin isolated from gestational pyelonephritis-associated *Escherichia coli* that binds to decay-accelerating factor. *Infect Immun* 65: 4309-4318.
- PICARD B, GARCIA JS, GOURIOU S, DURIEZ P, BRAHIMI N, BINGEN E, ELION J AND DENAMUR E. 1999. The link between phylogeny and virulence in *Escherichia coli* extraintestinal infection. *Infect Immun* 67: 546-553.
- PICARD B, JOURNET-MANCY C, PICARD-PASQUIER N AND GOULLET P. 1993. Genetic structures of the B2 and B1 *Escherichia coli* strains responsible for extraintestinal infections. *Journal of general microbiology* 139: 3079-3088.
- POURBAKHS SH SA, DHO-MOULIN M, BREE A, DESAUTELS C, MARTINEAU-DOIZE B AND FAIRBROTHER JM. 1997. Localization of the in vivo expression of P and F1 fimbriae in chickens experimentally inoculated with pathogenic *Escherichia coli*. *Microb Pathog* 22: 331-341.
- POUTTU R, PUUSTINEN T, VIRKOLA R, HACKER J, KLEMM P AND KORHONEN TK. 1999. Amino acid residue Ala-62 in the FimH fimbrial adhesin is critical for the adhesiveness of meningitis-associated *Escherichia coli* to collagens. *Mol Microbiol* 31: 1747-1757.
- PRASADARAO NV, WASS CA, HACKER J, JANN K AND KIM KS. 1993. Adhesion of S-fimbriated *Escherichia coli* to brain glycolipids mediated by sfaA gene-encoded protein of S-fimbriae. *J Biol Chem* 268: 10356-10363.
- PRATT LA AND KOLTER R. 1998. Genetic analysis of *Escherichia coli* biofilm formation: roles of flagella, motility, chemotaxis and type I pili. *Mol Microbiol* 30: 285-293.
- PROVENCE DL AND CURTISS R, 3RD. 1992. Role of *crl* in avian pathogenic *Escherichia coli*: a knockout mutation of *crl* does not affect hemagglutination activity, fibronectin binding, or Curli production. *Infect Immun* 60: 4460-4467.

- PROVENCE DL AND CURTISS R, 3RD. 1994. Isolation and characterization of a gene involved in hemagglutination by an avian pathogenic *Escherichia coli* strain. *Infect Immun* 62: 1369-1380.
- PRUSS BM, BESEMANN C, DENTON A AND WOLFE AJ. 2006. A complex transcription network controls the early stages of biofilm development by *Escherichia coli*. *J Bacteriol* 188: 3731-3739.
- RAYMOND KN, DERTZ EA AND KIM SS. 2003. Enterobactin: an archetype for microbial iron transport. *Proc Natl Acad Sci U S A* 100: 3584-3588.
- REIGSTAD CS, HULTGREN SJ AND GORDON JI. 2007. Functional genomic studies of uropathogenic *Escherichia coli* and host urothelial cells when intracellular bacterial communities are assembled. *J Biol Chem* 282: 21259-21267.
- REISNER A, HAAGENSEN JA, SCHEMBRI MA, ZECHNER EL AND MOLIN S. 2003. Development and maturation of *Escherichia coli* K-12 biofilms. *Mol Microbiol* 48: 933-946.
- REISNER A, KROGFELT KA, KLEIN BM, ZECHNER EL AND MOLIN S. 2006. In vitro biofilm formation of commensal and pathogenic *Escherichia coli* strains: impact of environmental and genetic factors. *J Bacteriol* 188: 3572-3581.
- RENNIE RP AND ARBUTHNOTT JP. 1974. Partial characterisation of *Escherichia coli* haemolysin. *J Med Microbiol* 7: 179-188.
- RIPPERE-LAMPE KE, O'BRIEN AD, CONRAN R AND LOCKMAN HA. 2001. Mutation of the gene encoding cytotoxic necrotizing factor type 1 (cnf(1)) attenuates the virulence of uropathogenic *Escherichia coli*. *Infect Immun* 69: 3954-3964.
- RIZVI M AND KUMAR S. 2003. Serum resistance of *Escherichia coli* strains causing urinary tract infection and diarrhoea in relation to alpha haemolysin production and O type. *Indian journal of pathology & microbiology* 46: 504-506.
- ROBERTS JA, KAACK MB, BASKIN G, CHAPMAN MR, HUNSTAD DA, PINKNER JS AND HULTGREN SJ. 2004. Antibody responses and protection from pyelonephritis following vaccination with purified *Escherichia coli* PapDG protein. *The Journal of urology* 171: 1682-1685.
- RODRIGUEZ-SIEK KE, GIDDINGS CW, DOETKOTT C, JOHNSON TJ, FAKHR MK AND NOLAN LK. 2005. Comparison of *Escherichia coli* isolates implicated in human urinary tract infection and avian colibacillosis. *Microbiology* 151: 2097-2110.
- ROMLING U. 2005. Characterization of the rdar morphotype, a multicellular behaviour in *Enterobacteriaceae*. *Cellular and molecular life sciences : CMLS* 62: 1234-1246.
- RON EZ. 2006. Host specificity of septicemic *Escherichia coli*: human and avian pathogens. *Current opinion in microbiology* 9: 28-32.
- RUSSO T, BROWN JJ, JODUSH ST AND JOHNSON JR. 1996. The O4 specific antigen moiety of lipopolysaccharide but not the K54 group 2 capsule is important for urovirulence of an extraintestinal isolate of *Escherichia coli*. *Infect Immun* 64: 2343-2348.
- RUSSO TA AND JOHNSON JR. 2000. Proposal for a new inclusive designation for extraintestinal pathogenic isolates of *Escherichia coli*: ExPEC. *The Journal of infectious diseases* 181: 1753-1754.
- RUSSO TA AND JOHNSON JR. 2003. Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem. *Microbes Infect* 5: 449-456.

- RUSSO TA, LIANG Y AND CROSS AS. 1994. The presence of K54 capsular polysaccharide increases the pathogenicity of *Escherichia coli* in vivo. *The Journal of infectious diseases* 169: 112-118.
- RUSSO TA, MCFADDEN CD, CARLINO-MACDONALD UB, BEANAN JM, BARNARD TJ AND JOHNSON JR. 2002. IroN functions as a siderophore receptor and is a urovirulence factor in an extraintestinal pathogenic isolate of *Escherichia coli*. *Infect Immun* 70: 7156-7160.
- RUSSO TA, MOFFITT MC, HAMMER CH AND FRANK MM. 1993. TnphoA-mediated disruption of K54 capsular polysaccharide genes in *Escherichia coli* confers serum sensitivity. *Infect Immun* 61: 3578-3582.
- SAHU BB, SUMIT R, SRIVASTAVA SK AND BHATTACHARYYA MK. 2012. Sequence based polymorphic (SBP) marker technology for targeted genomic regions: its application in generating a molecular map of the *Arabidopsis thaliana* genome. *BMC genomics* 13: 20.
- SALDANA Z, XICOHTENCATL-CORTES J, AVELINO F, PHILLIPS AD, KAPER JB, PUENTE JL AND GIRON JA. 2009. Synergistic role of curli and cellulose in cell adherence and biofilm formation of attaching and effacing *Escherichia coli* and identification of Fis as a negative regulator of curli. *Environmental microbiology* 11: 992-1006.
- SARKAR SF AND GUTTMAN DS. 2004. Evolution of the core genome of *Pseudomonas syringae*, a highly clonal, endemic plant pathogen. *Appl Environ Microbiol* 70: 1999-2012.
- SCHEMBRI MA AND KLEMM P. 2001a. Biofilm formation in a hydrodynamic environment by novel fimh variants and ramifications for virulence. *Infect Immun* 69: 1322-1328.
- SCHEMBRI MA AND KLEMM P. 2001b. Coordinate gene regulation by fimbriae-induced signal transduction. *The EMBO journal* 20: 3074-3081.
- SCHEMBRI MA, SOKURENKO EV AND KLEMM P. 2000. Functional flexibility of the FimH adhesin: insights from a random mutant library. *Infect Immun* 68: 2638-2646.
- SCHEMBRI MA, USSERY DW, WORKMAN C, HASMAN H AND KLEMM P. 2002. DNA microarray analysis of fim mutations in *Escherichia coli*. *Molecular genetics and genomics* : MGG 267: 721-729.
- SCHIERACK P, ROMER A, JORES J, KASPAR H, GUENTHER S, FILTER M, EICHBERG J AND WIELER LH. 2009. Isolation and characterization of intestinal *Escherichia coli* clones from wild boars in Germany. *Appl Environ Microbiol* 75: 695-702.
- SCHMOLL T, HOSCHUTZKY H, MORSCHHAUSER J, LOTTSPREICH F, JANN K AND HACKER J. 1989. Analysis of genes coding for the sialic acid-binding adhesin and two other minor fimbrial subunits of the S-fimbrial adhesin determinant of *Escherichia coli*. *Mol Microbiol* 3: 1735-1744.
- SCHOULER C, KOFFMANN F, AMORY C, LEROY-SETRIN S AND MOULIN-SCHOULEUR M. 2004. Genomic subtraction for the identification of putative new virulence factors of an avian pathogenic *Escherichia coli* strain of O2 serogroup. *Microbiology* 150: 2973-2984.
- SCHOULEUR M, REPERANT M, LAURENT S, BREE A, MIGNON-GRASTEAU S, GERMON P, RASSCHAERT D AND SCHOULER C. 2007. Extraintestinal pathogenic *Escherichia coli* strains of avian and human origin: link between phylogenetic relationships and common virulence patterns. *Journal of clinical microbiology* 45: 3366-3376.

- SCHOULEUR M, SCHOULER C, TAILLIEZ P, KAO MR, BREE A, GERMON P, OSWALD E, MAINIL J, BLANCO M AND BLANCO J. 2006. Common virulence factors and genetic relationships between O18:K1:H7 *Escherichia coli* isolates of human and avian origin. *Journal of clinical microbiology* 44: 3484-3492.
- SCHRAG S AND SCHUCHAT A. 2005. Prevention of neonatal sepsis. *Clinics in perinatology* 32: 601-615.
- SEGATA N AND HUTTENHOWER C. 2011. Toward an efficient method of identifying core genes for evolutionary and functional microbial phylogenies. *PLoS One* 6: e24704.
- SERVIN AL. 2005. Pathogenesis of Afa/Dr diffusely adhering *Escherichia coli*. *Clin Microbiol Rev* 18: 264-292.
- SHEIKH J, HICKS S, DALL'AGNOL M, PHILLIPS AD AND NATARO JP. 2001. Roles for Fis and YafK in biofilm formation by enteroaggregative *Escherichia coli*. *Mol Microbiol* 41: 983-997.
- SHELUDCHENKO MS, HUYGENS F AND HARGREAVES MH. 2010. Highly discriminatory single-nucleotide polymorphism interrogation of *Escherichia coli* by use of allele-specific real-time PCR and eBURST analysis. *Appl Environ Microbiol* 76: 4337-4345.
- SHELUDCHENKO MS, HUYGENS F AND HARGREAVES MH. 2011. Human-specific *E.coli* single nucleotide polymorphism (SNP) genotypes detected in a South East Queensland waterway, Australia. *Environmental science & technology* 45: 10331-10336.
- SHIMODAIRA H. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* 17: 1246-1247.
- SHIMODAIRA H. 2002. An approximately unbiased test of phylogenetic tree selection. *Syst Biol* 51: 492-508.
- SHIMODAIRA H AND HASEGAWA M. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* 17: 1246-1247.
- SHORT EC AND KURTZ HJ. 1971. Properties of the Hemolytic Activities of *Escherichia coli*. *Infect Immun* 3: 678-687.
- SIEGFRIED L, KMETOVA M, JANIGOVA V, SASINKA M AND TAKACOVA V. 1995. Serum response of *Escherichia coli* strains causing dyspepsia and urinary tract infection: relation to alpha-hemolysin production and O type. *Infect Immun* 63: 4543-4545.
- SIMMS AN AND MOBLEY HL. 2008. PapX, a P fimbrial operon-encoded inhibitor of motility in uropathogenic *Escherichia coli*. *Infect Immun* 76: 4833-4841.
- SINGH A, GOERING RV, SIMJEE S, FOLEY SL AND ZERVOS MJ. 2006. Application of molecular techniques to the study of hospital infection. *Clin Microbiol Rev* 19: 512-530.
- SKAAR EP. 2010. The battle for iron between bacterial pathogens and their vertebrate hosts. *PLoS pathogens* 6: e1000949.
- SKYBERG JA, JOHNSON TJ, JOHNSON JR, CLABOTS C, LOGUE CM AND NOLAN LK. 2006. Acquisition of avian pathogenic *Escherichia coli* plasmids by a commensal *E. coli* isolate enhances its abilities to kill chicken embryos, grow in human urine, and colonize the murine kidney. *Infect Immun* 74: 6287-6292.
- SKYBERG JA, SIEK KE, DOETKOTT C AND NOLAN LK. 2007. Biofilm formation by avian *Escherichia coli* in relation to media, source and phylogeny. *Journal of applied microbiology* 102: 548-554.
- SLOAN LM, DURESKO BJ, GUSTAFSON DR AND ROSENBLATT JE. 2008. Comparison of real-time PCR for detection of the tcdC gene with four toxin immunoassays and

- culture in diagnosis of *Clostridium difficile* infection. Journal of clinical microbiology 46: 1996-2001.
- SMITH. 2007. Iron metabolism at the host pathogen interface: lipocalin 2 and the pathogen-associated iroA gene cluster. Int J Biochem Cell Biol 39(10): 1776-1780
- SMITH JL, FRATAMICO PM AND GUNTHER NW. 2007. Extraintestinal pathogenic *Escherichia coli*. Foodborne pathogens and disease 4: 134-163.
- SMITH JM. 1992. Analyzing the mosaic structure of genes. J Mol Evol 34: 126-129.
- SNYDER JA, HAUGEN BJ, LOCKATELL CV, MARONCLE N, HAGAN EC, JOHNSON DE, WELCH RA AND MOBLEY HL. 2005. Coordinate expression of fimbriae in uropathogenic *Escherichia coli*. Infect Immun 73: 7588-7596.
- SODERBLOM T, LAESTADIUS A, OXHAMRE C, APERIA A AND RICHTER-DAHLFORS A. 2002. Toxin-induced calcium oscillations: a novel strategy to affect gene regulation in target cells. Int J Med Microbiol 291: 511-515.
- SOKURENKO EV, CHESNOKOVA V, DOYLE RJ AND HASTY DL. 1997. Diversity of the *Escherichia coli* type 1 fimbrial lectin. Differential binding to mannosides and uroepithelial cells. J Biol Chem 272: 17880-17886.
- SOKURENKO EV, CHESNOKOVA V, DYKHUIZEN DE, OFEK I, WU XR, KROGFELT KA, STRUVE C, SCHEMBRI MA AND HASTY DL. 1998. Pathogenic adaptation of *Escherichia coli* by natural variation of the FimH adhesin. Proc Natl Acad Sci U S A 95: 8922-8926.
- SOKURENKO EV, COURTNEY HS, ABRAHAM SN, KLEMM P AND HASTY DL. 1992. Functional heterogeneity of type 1 fimbriae of *Escherichia coli*. Infect Immun 60: 4709-4719.
- SOKURENKO EV, COURTNEY HS, MASLOW J, SIITONEN A AND HASTY DL. 1995. Quantitative differences in adhesiveness of type 1 fimbriated *Escherichia coli* due to structural differences in fimH genes. J Bacteriol 177: 3680-3686.
- SOKURENKO EV, COURTNEY HS, OHMAN DE, KLEMM P AND HASTY DL. 1994. FimH family of type 1 fimbrial adhesins: functional heterogeneity due to minor sequence variations among fimH genes. J Bacteriol 176: 748-755.
- SOKURENKO EV, FELDGARDEN M, TRINTCHINA E, WEISSMAN SJ, AVAGYAN S, CHATTOPADHYAY S, JOHNSON JR AND DYKHUIZEN DE. 2004. Selection footprint in the FimH adhesin shows pathoadaptive niche differentiation in *Escherichia coli*. Mol Biol Evol 21: 1373-1383.
- SOKURENKO EV, HASTY DL AND DYKHUIZEN DE. 1999. Pathoadaptive mutations: gene loss and variation in bacterial pathogens. Trends Microbiol 7: 191-195.
- SOLOAGA A, VEIGA MP, GARCIA-SEGURA LM, OSTOLAZA H, BRASSEUR R AND GONI FM. 1999. Insertion of *Escherichia coli* alpha-haemolysin in lipid bilayers as a non-transmembrane integral protein: prediction and experiment. Mol Microbiol 31: 1013-1024.
- SOTO GE AND HULTGREN SJ. 1999. Bacterial adhesins: common themes and variations in architecture and assembly. J Bacteriol 181: 1059-1071.
- SOTO SM, SMITHSON A, HORCAJADA JP, MARTINEZ JA, MENSA JP AND VILA J. 2006. Implication of biofilm formation in the persistence of urinary tract infection caused by uropathogenic *Escherichia coli*. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases 12: 1034-1036.
- SOTO SM, SMITHSON A, MARTINEZ JA, HORCAJADA JP, MENSA J AND VILA J. 2007. Biofilm formation in uropathogenic *Escherichia coli* strains: relationship with

- prostatitis, urovirulence factors and antimicrobial resistance. *The Journal of urology* 177: 365-368.
- SPIGAGLIA P, CARATTOLI A, BARBANTI F AND MASTRANTONIO P. 2010. Detection of *gyrA* and *gyrB* mutations in *Clostridium difficile* isolates by real-time PCR. *Mol Cell Probes* 24: 61-67.
- STAWSKI G, NIELSEN L, ORSKOV F AND ORSKOV I. 1990. Serum sensitivity of a diversity of *Escherichia coli* antigenic reference strains. Correlation with an LPS variation phenomenon. *APMIS : acta pathologica, microbiologica, et immunologica Scandinavica* 98: 828-838.
- STEHLING EG CT, FERREIRA A, SILVEIRA WD. 2003. Adhesion and invasion characteristics of a Septicaemic avian *Escherichia coli* strain are plasmid mediated. *J Appl Res Vet Med* 1: 27-36.
- STENUTZ R, WEINTRAUB A AND WIDMALM G. 2006. The structures of *Escherichia coli* O-polysaccharide antigens. *FEMS microbiology reviews* 30: 382-403.
- STOLL ET AL HN, HIGGINS RD, FANAROFF AA, DUARA S, GOLDBERG R, LAPTOOK A, WALSH M, OH W, HALE E. 2005. Very low birth weight preterm infants with early onset neonatal sepsis: the predominance of gram-negative infections continues in the National Institute of Child Health and Human Development Neonatal Research Network, 2002-2003. *Pediatr Infect Dis J* 2005 Jul; 24: 635-639.
- SUERBAUM S, FRIEDRICH S, LEYING H AND OPFERKUCH W. 1994. Expression of capsular polysaccharide determines serum resistance in *Escherichia coli* K92. *Zentralblatt fur Bakteriologie : international journal of medical microbiology* 281: 146-157.
- SUSKO E, LEIGH J, DOOLITTLE WF AND BAPTESTE E. 2006. Visualizing and assessing phylogenetic congruence of core gene sets: a case study of the gamma-proteobacteria. *Mol Biol Evol* 23: 1019-1030.
- SVANBORG-EDEN C, HAGBERG L, HULL R, HULL S, MAGNUSSON KE AND OHMAN L. 1987. Bacterial virulence versus host resistance in the urinary tracts of mice. *Infect Immun* 55: 1224-1232.
- SVANBORG C ET AL. 2001. The 'innate' host response protects and damages the infected urinary tract. *Annals of medicine* 33: 563-570.
- SYVANEN AC. 2001. Accessing genetic variation: genotyping single nucleotide polymorphisms. *Nature reviews Genetics* 2: 930-942.
- TAMURA K, PETERSON D, PETERSON N, STECHER G, NEI M AND KUMAR S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28: 2731-2739.
- TARTOF SY, SOLBERG OD AND RILEY LW. 2007. Genotypic analyses of uropathogenic *Escherichia coli* based on *fimH* single nucleotide polymorphisms (SNPs). *J Med Microbiol* 56: 1363-1369.
- TAYLOR PW. 1983. Bactericidal and bacteriolytic activity of serum against gram-negative bacteria. *Microbiological reviews* 47: 46-83.
- TAYLOR PW AND HUGHES C. 1978. Plasmid carriage and the serum sensitivity of enterobacteria. *Infect Immun* 22: 10-17.
- TENAILLON O, SKURNIK D, PICARD B AND DENAMUR E. 2010. The population genetics of commensal *Escherichia coli*. *Nat Rev Microbiol* 8: 207-217.

- TEWARI R, IKEDA T, MALAVIYA R, MACGREGOR JI, LITTLE JR, HULTGREN SJ AND ABRAHAM SN. 1994. The PapG tip adhesin of P fimbriae protects *Escherichia coli* from neutrophil bactericidal activity. *Infect Immun* 62: 5296-5304.
- THANASSI DG, SAULINO ET, LOMBARDO MJ, ROTH R, HEUSER J AND HULTGREN SJ. 1998. The PapC usher forms an oligomeric channel: implications for pilus biogenesis across the outer membrane. *Proc Natl Acad Sci U S A* 95: 3146-3151.
- TIVENDALE KA, ALLEN JL AND BROWNING GF. 2009. Plasmid-borne virulence-associated genes have a conserved organization in virulent strains of avian pathogenic *Escherichia coli*. *Journal of clinical microbiology* 47: 2513-2519.
- TIVENDALE KA, LOGUE CM, KARIYAWASAM S, JORDAN D, HUSSEIN A, LI G, WANNEMUEHLER Y AND NOLAN LK. 2010. Avian-pathogenic *Escherichia coli* strains are similar to neonatal meningitis *E. coli* strains and are able to cause meningitis in the rat model of human disease. *Infect Immun* 78: 3412-3419.
- TORRES AG, REDFORD P, WELCH RA AND PAYNE SM. 2001. TonB-dependent systems of uropathogenic *Escherichia coli*: aerobactin and heme transport and TonB are required for virulence in the mouse. *Infect Immun* 69: 6179-6185.
- TOUCHON M ET AL. 2009. Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS genetics* 5: e1000344.
- UHLEN P, LAESTADIUS A, JAHNUKAINEN T, SODERBLOM T, BACKHED F, CELSI G, BRISMAR H, NORMARK S, APERIA A AND RICHTER-DAHLFORS A. 2000. Alpha-haemolysin of uropathogenic *E. coli* induces Ca²⁺ oscillations in renal epithelial cells. *Nature* 405: 694-697.
- UHLICH GA, KEEN JE AND ELDER RO. 2001. Mutations in the *csgD* promoter associated with variations in curli expression in certain strains of *Escherichia coli* O157:H7. *Appl Environ Microbiol* 67: 2367-2370.
- ULETT GC, MABBETT AN, FUNG KC, WEBB RI AND SCHEMBRI MA. 2007. The role of F9 fimbriae of uropathogenic *Escherichia coli* in biofilm formation. *Microbiology* 153: 2321-2331.
- UNHANAND M, MUSTAFA MM, MCCRACKEN GH, JR. AND NELSON JD. 1993. Gram-negative enteric bacillary meningitis: a twenty-one-year experience. *The Journal of pediatrics* 122: 15-21.
- USEIN CR, DAMIAN M, TATU-CHITOIU D, CAPUSA C, FAGARAS R, TUDORACHE D, NICA M AND LE BOUGUENEC C. 2001. Prevalence of virulence genes in *Escherichia coli* strains isolated from Romanian adult urinary tract infection cases. *Journal of cellular and molecular medicine* 5: 303-310.
- WAGNER W, KUHN M AND GOEBEL W. 1988. Active and inactive forms of hemolysin (HlyA) from *Escherichia coli*. *Biological chemistry Hoppe-Seyler* 369: 39-46.
- WANG MC, TSENG CC, WU AB, HUANG JJ, SHEU BS AND WU JJ. 2009. Different roles of host and bacterial factors in *Escherichia coli* extra-intestinal infections. *Clin Microbiol Infec* 15: 372-379.
- WANG S, NIU C, SHI Z, XIA Y, YAQOUB M, DAI J AND LU C. 2011. Effects of *ibeA* deletion on virulence and biofilm formation of avian pathogenic *Escherichia coli*. *Infect Immun* 79: 279-287.
- WANG Y, TANG C, YU X, XIA M AND YUE H. 2010. Distribution of serotypes and virulence-associated genes in pathogenic *Escherichia coli* isolated from ducks. *Avian pathology : journal of the WVPA* 39: 297-302.

- WEBER JL AND MAY PE. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American journal of human genetics* 44: 388-396.
- WEISER JN AND GOTSCHLICH EC. 1991. Outer membrane protein A (OmpA) contributes to serum resistance and pathogenicity of *Escherichia coli* K-1. *Infect Immun* 59: 2252-2258.
- WEISSMAN SJ, BESKHLEBNAYA V, CHESNOKOVA V, CHATTOPADHYAY S, STAMM WE, HOOTON TM AND SOKURENKO EV. 2007. Differential stability and trade-off effects of pathoadaptive mutations in the *Escherichia coli* FimH adhesin. *Infect Immun* 75: 3548-3555.
- WEISSMAN SJ, CHATTOPADHYAY S, APRIKIAN P, OBATA-YASUOKA M, YAROVA-YAROVAYA Y, STAPLETON A, BA-THEIN W, DYKHUIZEN D, JOHNSON JR AND SOKURENKO EV. 2006. Clonal analysis reveals high rate of structural mutations in fimbrial adhesins of extraintestinal pathogenic *Escherichia coli*. *Mol Microbiol* 59: 975-988.
- WEISSMAN SJ ET AL. 2012. High-resolution two-locus clonal typing of extraintestinal pathogenic *Escherichia coli*. *Appl Environ Microbiol* 78: 1353-1360.
- WILES TJ, KULESUS RR AND MULVEY MA. 2008. Origins and virulence mechanisms of uropathogenic *Escherichia coli*. *Experimental and molecular pathology* 85: 11-19.
- WILLIAMS JG, KUBELIK AR, LIVAK KJ, RAFALSKI JA AND TINGEY SV. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res* 18: 6531-6535.
- WIRTH T ET AL. 2006. Sex and virulence in *Escherichia coli*: an evolutionary perspective. *Mol Microbiol* 60: 1136-1151.
- WIZEMANN TM, ADAMOU JE AND LANGERMANN S. 1999. Adhesins as targets for vaccine development. *Emerging infectious diseases* 5: 395-403.
- WOLFF D, BRUNING T AND GERRITZEN A. 2009. Rapid detection of the *Clostridium difficile* ribotype 027 tcdC gene frame shift mutation at position 117 by real-time PCR and melt curve analysis. *European journal of clinical microbiology & infectious diseases* : official publication of the European Society of Clinical Microbiology 28: 959-962.
- WOOLEY RE, NOLAN LK, BROWN J, GIBBS PS, GIDDINGS CW AND TURNER KS. 1993. Association of K-1 capsule, smooth lipopolysaccharides, traT gene, and Colicin V production with complement resistance and virulence of avian *Escherichia coli*. *Avian diseases* 37: 1092-1096.
- WRIGHT KJ AND HULTGREN SJ. 2006. Sticky fibers and uropathogenesis: bacterial adhesins in the urinary tract. *Future microbiology* 1: 75-87.
- WRIGHT KJ, SEED PC AND HULTGREN SJ. 2005. Uropathogenic *Escherichia coli* flagella aid in efficient urinary tract colonization. *Infect Immun* 73: 7657-7668.
- WU G, EHRLICH R, MAFURA M, STOKES M, SMITH N, PRITCHARD GC AND WOODWARD MJ. 2012. *Escherichia coli* isolates from extraintestinal organs of livestock animals harbour diverse virulence genes and belong to multiple genetic lineages. *Vet Microbiol* 160: 197-206.
- WU XR, SUN TT AND MEDINA JJ. 1996. In vitro binding of type 1-fimbriated *Escherichia coli* to uroplakins Ia and Ib: relation to urinary tract infections. *Proc Natl Acad Sci U S A* 93: 9630-9635.
- WULLT B, BERGSTEN G, CONNELL H, ROLLANO P, GEBRETSADIK N, HULL R AND SVANBORG C. 2000. P fimbriae enhance the early establishment of *Escherichia coli* in the human urinary tract. *Mol Microbiol* 38: 456-464.

- XIA Y, GALLY D, FORSMAN-SEMB K AND UHLIN BE. 2000. Regulatory cross-talk between adhesin operons in *Escherichia coli*: inhibition of type 1 fimbriae expression by the PapB protein. *The EMBO journal* 19: 1450-1457.
- XIE Y, KIM KJ AND KIM KS. 2004. Current concepts on *Escherichia coli* K1 translocation of the blood-brain barrier. *FEMS immunology and medical microbiology* 42: 271-279.
- YANG Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci* 13: 555-556.
- ZIMECKI M, CHODACZEK G, KOCIEBA M AND KRUZEL ML. 2004. Lethality in LPS-induced endotoxemia in C3H/HeCr mice is associated with prevalence of proinflammatory cytokines: lack of protective action of lactoferrin. *FEMS immunology and medical microbiology* 42: 167-172.
- ZOGAJ X, NIMTZ M, ROHDE M, BOKRANZ W AND ROMLING U. 2001. The multicellular morphotypes of *Salmonella typhimurium* and *Escherichia coli* produce cellulose as the second component of the extracellular matrix. *Mol Microbiol* 39: 1452-1463.