CHAPTER 1

INTRODUCTION

Neonatal sepsis is clinically defined as the presence of systemic signs and symptoms of infection with bacteriological isolation of an infectious agent from blood or cerebrospinal fluid (CSF) in the neonatal period i.e. first four weeks of life (Paolucci *et al*., 2011; Campbell *et al*., 2011). Sepsis in the newborns is a challenging problem, accounting for 30-50% of total neonatal deaths each year in the developing countries (Black *et al*., 2010; Das *et al*., 2011; Tripathi and Malik, 2010). Approximately, 20% of neonates develop sepsis with about 1% death due to sepsis related causes (Stoll, 1997). In India, sepsis and pneumonia together accounts for being the commonest cause of neonatal deaths, responsible for 30.4% of total neonatal deaths (Lahariya C and Paul VK, 2010). Early and accurate diagnosis is the most crucial step in preventing the disease. Several limitations in timely detection of the disease, resource-poor settings, neonatal risk factors and susceptibility towards an infection, a huge repertoire of causative organisms and multidrug resistance all make this disease daunting.

On the basis of onset of symptoms, neonatal sepsis is broadly classified into two types, early onset sepsis (EOS) and late onset sepsis (LOS) (Bizzarro *et al*., 2005). EOS presents within 72 hrs of life and is caused by vertical transmission of microorganisms from mother before and during delivery (Kaftan and Kinney, 1998). The organisms are horizontally transmitted (nosocomial or community acquired) in LOS which occurs after 72 hrs of life (Bizzarro *et al*., 2005, Shah and Padbury, 2014). This classification has significant contribution in understanding the mode of transmission, the causative organisms for the disease and the expected outcomes of infection which actually guides in rapid and proper diagnosis and treatment of the disease (Shah and Padbury, 2014; Paolucci *et al*., 2011). The severity of illness is more in EOS than LOS. The neonate is at high risk of mortality and morbidity in

EOS, whereas the frequency of LOS is more in low-income countries. The rigorous practices of home deliveries in a completely unhygienic and unclean environment, poor intrapartum and postnatal infection-control practices, overcrowding and understaffing of neonatal intensive care unit (NICU) and lack of knowledge about the importance of proper hygiene practices among staffs and workers of NICU are the predisposing factors for the increased frequency of LOS in developing countries (Edmond and Zaidi, 2010; Zaidi *et al*., 2005; Pittet and Boyce, 2001).

A wide range of organisms are implicated in sepsis, of which Gram negatives are more common in developing countries. Gram negative organisms are responsible for 18-78% of all neonatal sepsis (Mutlu *et al*., 2011; Macharashvili *et al*., 2009; Kristof *et al*., 2009). The risk of death is higher in Gram negative sepsis. Among the broad spectrum of Gram negative pathogens involved in sepsis of newborns, *Escherichia coli* occupies a pivotal position in developing countries, specifically in South Asia and Africa (Sanghvi and Tudehope, 1996; Stoll *et al*., 2002). *E. coli* has emerged as the leading cause of neonatal sepsis in preterm infants and the second most common pathogen in term infants (Stoll *et al*., 2011). *E. coli* is the major pathogen for sepsis related mortality among very low birth weight (VLBW) infants (24.5%) (Weston *et al*., 2011). *E. coli* is frequently associated with incidence of EOS as it can colonize mother's genital tract and infect neonates at the time of delivery. Hence, *E. coli* is now considered to be the most important agent for EONS (Lin *et al*., 2011; Stoll *et al*., 2011). In addition, cases of poor outcomes leading to high mortality in newborns with early-onset *E. coli* sepsis is also evidenced (Mayor-Lynn *et al*., 2005). Though generally implicated in EOS, *E. coli* has also been found to be associated with LOS in developing countries in recent years (Viswanathan, 2012).

E. coli isolates, on the basis of both genetic and clinical characteristics, are broadly classified into 3 major groups: commensal *E. coli*, intestinal pathogenic *E. coli* (IPEC) and extraintestinal pathogenic *E. coli* (ExPEC). Commensal *E. coli* strains represent a major portion of normal facultative intestinal flora which serenely coexist with the host and do not cause disease within the host intestinal tract (Russo and Johnson, 2000). Certain commensal *E. coli* isolates can exit gut and can cause diseases in different extraintestinal tissues, specifically in immunocompromised hosts. These are termed as ExPECs which cause several extraintestinal diseases such as urinary tract infection (UTI), pneumonia, intra-abdominal infections, (like cystitis, pyelonephritis), osteomyelitis, cellulitis, soft-tissue infection, meningitis and sepsis. ExPECs are unable to cause gastrointestinal diseases like the IPECs (Johnson and Russo, 2002, Russo and Johnson, 2003). However, they can asymptomatically colonize the host intestinal tract for prolonged period of time and under certain circumstances can exit gut and enter into sterile body sites to cause diverse extraintestinal diseases (Johnson and Russo, 2001).

E. coli isolates are classified into four major phylogenetic groups namely 'A', 'B1', 'B2' and 'D'(Herzer *et al*., 1990; Clermont *et al*., 2000). *E. coli* isolates causing extra-intestinal infections are mainly derived from phylogroup B2 and to a lesser extent phylogroup D (Salender *et al*., 1986; Bingen *et al*., 1998; Picard *et al*., 1999; Johnson and Stell, 2000) whereas most commensal isolates belong to phylogroups A and B1 (Picard *et al*., 1999; Duriez *et al*., 2001; Das *et al*., 2013). However, there are a few exceptions which clearly demonstrate the existence of the non-pathogenic phylogroups A and B1 among the ExPEC isolates (Abdallah *et al*., 2011). Due to their consistent associations with various extraintestinal diseases, groups B2 and D consist

of diverse evolutionary lineages which are regarded as virulent clones, as usually defined based on O:K:H serotypes (Johnson and Stell, 2000).

A hallmark of such virulent *E. coli* clones is their possession of specialized virulence factor (VF) genes (distinct from those of the IPEC) which enable them to cause infections outside the gastrointestinal tract, in both normal and compromised hosts (Picard *et al*., 1999; Johnson and Stell, 2000). These virulence attributes are generally absent in fecal isolates (Watt *et al*., 2003; Ron, 2010; Chapman *et al*., 2006; Mokady *et al*., 2005). The characteristic virulence traits that are present in the septicemic isolates include various adhesins (e.g. type S and type P pili), factors for iron acquisition (e.g. IucC, IroN), factors for invasion (IbeA), factors to subvert host defence systems (e.g. capsule, lipopolysaccharide) and toxins [like cytotoxic necrotizing factor (Cnf), alpha-hemolysin (Hly)]. These VFs essentially contribute in colonization and invasion of the host, disruption of host defence mechanisms, injury to host issues and stimulation of a deleterious host inflammatory response. In addition, pathogenicity-associated islands (PAIs) which comprise blocks of VF genes and provide a mechanism for coordinate horizontal transfer of VF genes between lineages and even between species are also considered as a unifying characteristic of ExPECs (Boyd and Hartl, 1998; Guyer *et al*., 1998; Kao *et al*., 1997). A study by Tullus *et al*. has emphasised the importance of considering host factors while evaluating possible VFs in pathogenic bacteria (Tullus *et al*., 1992) as host compromise can increase the pathogenic potential of certain VFs (Johnson, 1991; Johnson *et al*., 1988).

An addition to the list of VFs with assigned or assumed virulence functions are the extracellular secreted proteases of pathogens. Microbial proteases are one of the most critical and essential weapons of pathogenic bacteria to exacerbate the virulence of a

particular pathogen in the progression of a disease. There are several well known bacterial proteases that specifically interact with their hosts during a pathogenic infection and act on various extracellular or intracellular targets. The highly lethal Anthrax Toxin of *Bacillus anthracis* is a trimeric complex of three different proteins of which one is called the lethal factor. The lethal factor of the anthrax toxin is a metalloprotease that is able to specifically cleave and inactive MAP kinase kinases (Young and Collier, 2007). Botulinum neurotoxin (BoNT), the highly lethal toxin of *Clostridium botulinum* possesses a metalloprotease domain which is able to block acetylcholine release at peripheral nerve ending by the cleavage the SNAP-25 protein that plays a role in the storage and depletion of acetylcholine (Dolly and Aoki, 2006). *P. aeruginosa* also secretes several proteases involved in the pathogenesis of the organism, elastase B of *P. aeruginosa* cleaves bovine and human elastin and is responsible for the decreased levels of lung elastins in cystic fibrosis patients (Bruce *et al*., 1985). Elastase B can also degrade important biopolymer collagen and also secretory IgA (Heck *et al*., 1986; Heck *et al*., 1990).

Proteases can be catalytically classified into 6 different types depending on the active residues involved in catalysis. These include metalloproteases, serine proteases, aspartic proteases and cysteine proteases. The metallo, serine and aspartic proteases are the most commonly studied microbial proteases (Hase and Finkelstein, 1993).

One such group of serine proteases are SPATEs (Serine Protease Autotransporters of *Enterobacteriaceae*), which also belongs to the subfamily of autotransporters (ATs). ATs are a family of proteins that mediate their own secretion through the outer membrane of Gram negative bacteria. The secretion mechanism of ATs is known as type V secretion. The ATs are quite diverse in their function and can act as proteases, esterases, adhesins and lipases. SPATEs the subfamily of ATs are a group of secreted

serine proteases of *Enterobacteriaceae*. The member proteases of the family SPATEs share conserved identical structure comprising of three domains: an N-terminal signal sequence, a passenger domain containing the consensus GDSGSP serine protease motif and a C-terminal translocator β-domain. The N-terminal signal peptide is needed for targeting and export through the inner membrane, the C-terminal domain forms a β–barrel structure through the outer membrane for the translocation of the passenger domain to the extracellular milieu. The passenger domain is the functional domain of SPATEs and is diverse among member proteases of SPATE family. The passenger domain consists of an N-terminal globular domain harbouring a characteristic GDSGSP serine protease motif. The first serine of this motif forms a catalytic triad with a histidine and asparagine residue forming a substrate binding pocket. The C-terminal of the passenger domain folds into a β-helical stalk and is supposed to confer structural stability to the secreted protein. After translocation through the outer membrane, the passenger domains of all SPATEs are autoproteolyticaly cleaved at an FxxEVNNLNK motif in the linker domain which is located between the passenger domain and the β-domain (Henderson and Nataro, 2001; Bernstein, 2007; Dautin, 2010). SPATEs comprise more than 20 secreted proteases from *Enterobacteriaceae (Escherichia, Citrobacter, Shigella* and *Salmonella sp*.) (Yen *et al*., 2008). SPATEs are usually among the most abundant proteins secreted by the parental strain (Dautin, 2010).

Phylogenetically, SPATEs are classified into two groups, class I and class II based on the presence or absence of two smaller domains in the globular N-terminus of the passenger domain, designated as domain 3 and domain 4.Domain 3 possess a pair of cysteine residues in case of Class I SPATEs, but not in Class II (Ruiz-Perez and Nataro, 2014). Class I SPATEs are cytotoxic, whereas Class II SPATEs possess

domain 2 and are cytopathic having various intra and extracellular proteolytic targets. Members of Class I SPATEs include Pet (plasmid encoded toxin), Sat (secreted autotransporter toxin), SigA (Shigella IgA 1-protease like), EspP (extracellular serine protease plasmid (p0157)-encoded), EspC (EPEC secreted protein C) and Vat (vacuolating autotransporter toxin) from *E. coli* and *Shigella flexneri*. Class II SPATEs comprise of Pic (protease involved in intestinal colonization), SepA (*Shigella* extracellular protein A), EatA (ETEC autotransporter A) and Tsh (temperature-sensitive hemagglutinin) which has also been named as Hbp (Hemoglobin protease) (Dautin, 2010; Dutta, 2002; Boisen, 2009).

Although some recent studies have identified SPATEs in non-pathogenic isolates (Grozdanov *et al*., 2004; Sandt and Hill, 2000), they seem to be strongly associated with pathogenic isolates (Dautin, 2010; Parham *et al*., 2005; Restieri *et al*., 2007). Though SPATEs share common structural features, they have distinct substrate specificities and diverse mode of actions (Dautin, 2010; Dutta *et al*., 2002).

The distribution of SPATEs has been demonstrated in many studies among the diarrheagenic *E. coli.* In a study by Boison *et al*., the distribution of different subtypes of SPATEs has been reported among clinical isolates of enteroaggregative *E. coli* (EAEC), associated with several clinical conditions including pediatric diarrhea, travelers' diarrhea, persistent diarrhea among human immunodeficiency virus-infected patients. Both class I and class II SPATEs were found almost exclusively among the EAEC isolates in this study. In addition, the presence of SPATEs was also evidenced among clinical isolates of *Shigella* in this study (Boison *et al*., 2009). Another study by Restieri *et al*., has described the high prevalence of SPATEs among diarrheagenic *E. coli* isolates (Restieri *et al*., 2007).Furthermore, the significantly higher occurrence of Sat, a vacuolating toxin, was reported among clinical isolates of diffusely adherent

E. coli (DAEC), which are known to be involved in diarrhea among children of 4-5 years of age (Taddei *et al*., 2003).

The distribution of SPATEs is also well studied among various categories of ExPEC isolates, specifically among isolates causing UTI. The two most exclusively associated SPATEs with extraintestinal diseases were found to be Vat and Sat. Vat was found to be highly associated with uropathogenic *E. coli* (UPEC) isolates compared to the fecal isolates and is expressed during infection (Heimer *et al*., 2004). UPEC isolates having *vat* were recognized as more efficient colonizers of the urinary tract than those without *vat* and *vat* was identified among one of the essential marker genes for rapid diagnosis of UPECs by a simple multiplex polymerase chain reaction (PCR) (Spurbeck *et al*., 2012). The higher prevalence of SPATEs was also reported in a number of other studies among the various pathotypes of ExPEC isolates causing UTI, cystitis, pyelonephritis (Restieri *et al*., 2007; Parham *et al*., 2005; Messai *et al*., 2011). However, no study has been undertaken to illustrate the distribution of SPATEs among the *E. coli* isolates causing neonatal septicaemia.

In this study, septicemic, fecal and environmental *E. coli* isolates of different phylogroups were compared for the presence of different VFs and subtypes of SPATEs to determine the association of SPATEs with neonatal septicaemia. To the best of our knowledge, this study is the first to report the phylogenetic distribution of 10 different subtypes of SPATEs among clinical isolates of neonatal septicemic *E. coli* (NSEC).

In an attempt to purify, identify and characterize a secreted protease from a NSEC isolate, we identified YghJ as a secreted metalloprotease from a NSEC isolate. We further cloned, expressed and purified the metalloprotease YghJ and also studied the distribution of YghJ among neonatal septicemic and fecal *E. coli* isolates. Moreover, a

comparative analysis of the expression and secretion of YghJ between these two groups of isolates was also performed.

YghJ belongs to a large and diverse family of eukaryotic and prokaryotic proteins containing putative M60 metalloprotease domain. This domain is shared by a range of host-associated bacterial and eukaryotic microbes including commensals and pathogens of invertebrates and vertebrates. (Nakjang *et al*., 2012). YghJ has also very recently been identified to be secreted from Enterotoxigenic *E. coli* (ETEC) and shown to have mucinase activity. In another recent study, the chromosomally located *yghJ* gene has been documented to be secreted by 89% of the ETEC isolates, distributed over a diverse phylogenetic lineage and shown to be expressed during the course of infection of ETEC. This study suggests that YghJ is one of such proteins that trigger immune response to ETEC infection (Roy *et al*., 2010). Interestingly, YghJ was identified as one of the potential vaccine candidates for ExPEC isolates. The distribution study of YghJ showed that YghJ is distributed among all pathotypes of *E. coli* and also among commensal *E. coli* isolates. However, the incidence was higher in case of intestinal and extraintestinal pathogens compared to the commensals. Furthermore, the analysis of expression and secretion of YghJ revealed that YghJ is expressed but not secreted among the commensal isolates (Moriela *et al*., 2010). This study further suggested that YghJ is not only a mere surface antigen, but also might be associated with the virulence of diverse ExPEC isolates. Keeping this in mind, we have attempted to elucidate the mode of pathogenesis of YghJ secreted by NSEC isolates in neonatal septicaemia.

Sepsis is a complex pathophysiological process which can be characterized as an exacerbated systemic inflammatory response against an infection that involves the

overproduction of endogenous mediators. Cytokines are small protein mediators of low molecular weight and are synthesized in a regulated fashion by the cells of host's immune system during an infection process. The secreted proinflammatory cytokines then activate host defence by several mechanisms. These include recruitment of phagocytes and lymphocytes at the site of infection, activates microbiocidal potential of neutrophils and macrophages mainly through production of reactive oxygen and nitrogen species and potentiates phagocytosis. However, an amplified host response and overwhelming production of these cytokines can lead to endothelial dysfunction, characterized by vasodilation, increased vascular permeability, hypotension, multiple organ failure and ultimately shock and death. Proinflammatory cytokines released by numerous immune cells have been attributed to have key role in the pathogenesis of sepsis (Netea *et al*., 2003; Sikora *et al*., 2001; Schulte *et al*., 2013). Our study is the first to show that YghJ stimulates the production of several proinflammatory cytokines like IL-1α, IL-1β, IL-8 and TNF-α which are essentially implicated in the sepsis of newborns.

E. coli is amongst the major pathogens involved in neonatal sepsis, specifically in developing countries. This study was initially carried out to determine the distribution of different subtypes of SPATEs among different phylogroups of clinical isolates of *E. coli* causing neonatal septicaemia and to compare the prevalence of SPATEs among fecal and environmental *E. coli* isolates. This study is the first to report the significantly higher prevalence SPATEs among NSEC isolates compared to the fecal or environmental isolates. Vat and Sat was found to be the most predominant SPATEs. Interestingly, the occurrence of SPATEs was also higher compared to the other VFs tested implicating SPATEs as the most discriminatory trait studied here for NSEC isolates. The higher prevalence of SPATEs specifically among the non-

pathogenic phylogroups A and B1 gave us a hint of the association of SPATEs with NSEC isolates, the validity of which was further reinforced in the suckling mouse experiment. The animal experiment consolidated the pathogenic potential of the septicemic isolates belonging to phylogroups A and B1, possessing SPATEs but no other virulence determinants except *iucC* which is generally considered as a defensive VF and not directly linked to the causation of the disease. In addition, we identified an extracellular metalloprotease YghJ from a NSEC isolate which has previously been reported to be secreted from ETEC strains and having mucinase activity (Luo *et al*., 2014). Our study is the first to purify, identify YghJ from a NSEC isolate and to show that YghJ is able to trigger the production of proinflammatory cytokines and downregulates the induction of anti-inflmmatory cytokine IL-10. Furthermore, we illustrates that YghJ is expressed and secreted at a significantly higher rate in the septicemic isolates compared to the fecal *E. coli* isolates. In addition, we have also shown that YghJ is exclusively associated with the NSEC isolates harbouring the most prevalent SPATE, Vat. Hence, our study identifies a factor other than LPS responsible for the stimulation of proinflammatory cytokines which have been attributed to have a crucial role in the pathogenesis of sepsis.