SCUOLA DOTTORALE IN BIOLOGIA

SEZIONE SCIENZE BIOMOLECOLARI E CELLULARI

CICLO XXV

Genetic analysis of the antivirulence loci involved in the biosynthesis of nicotinamide adenine dinucleotide in enteroinvasive *Escherichia coli* and *Shigella.*

Dottorando

ROSA FIORAVANTI

Docente Guida: Prof.ssa MARIASSUNTA CASALINO

Coordinatore: Prof. PAOLO MARIOTTINI

INDEX

ABSTRACT

Shigella species are the primary etiologic agents of bacillary dysentery or shigellosis, which remains a significant threat to public health, particularly in less developed countries where sanitation is poor. There are an estimated 160 million shigellosis episodes worldwide yearly, with 1.1 million deaths, predominantly in children younger than 5 years of age. The genus *Shigella* is divided into four species: *S. dysenteriae, S. flexneri, S. boydii, and S. sonnei*. However, *Shigella* has been known for a long time to be closely related to *E. coli*. There are very few characteristics that can distinguish *Shigella* strains from enteroinvasive *E. coli* (EIEC), which also cause dysentery. Furthermore, recent phylogenetic studies have suggested that *Shigella* and EIEC form a single pathotype of *E. coli*. My PhD work will be mainly focused on the study of evolution of the *Shigella/*EIEC pathotype from an ancestor commensal *E. coli*. The critical event in the transition toward a pathogenic life-style has probably relied on the acquisition, likely through horizontal transfer, of the virulence plasmid (pINV), which encodes all the genes required for invasion and for intra- and intercellular spreading, including their positive activators. Together with the acquisition of the pINV, the *Shigella*/EIEC pathotype had lost some traits that were important for survival in the environment, but redundant for the life inside the host. The loss of these specific functions has been hypothesized to be functional to the pathogenetic lifestyle of the *Shigella*/EIEC because the expression of these genes, defined antivirulence loci (AVL), might have been detrimental for the expression of the newly acquired virulence genes or because might have been redundant in the new niches. Unlike most *E. coli* strains, *Shigella* strains require nicotinic acid supplementation for growth on minimal medium. The nicotinic acid requirement is due to mutation in two unlinked loci, *nadA* and *nadB,* encoding the enzyme complex that converts L-aspartate to quinolinate (QUIN), a precursor of the pathway leading to *de novo* synthesis of NAD. Recently has been reported that QUIN is a strong and specific inhibitor of several virulence phenotypes of *Shigella*. In this study, has been used a combination of genetic, molecular, and genomic approaches to analyse if the requirement for nicotinic acid may be selected during the evolution toward a pathogenic lifestyle in both *Shigella* and EIEC strains.

Genetic analysis of the antivirulence loci involved in the biosynthesis of NAD in enteroinvasive *Escherichia coli* and *Shigella.* The analysis of the evolutionary relationship among *Shigella* and *E. coli* indicates that *Shigella* strains have been derived repeatedly from different branches of the *E. coli* tree by convergent evolution involving both, gain and loss of genes. *Shigella* appears to have diverged from commensal *E.coli* more than enteroinvasive *E. coli*, a group of diarrhogenic *E .coli* which shares with *Shigella* the same

pathogenicity process. A characteristic of *Shigella* is their strict requirement for an exogenous source of nicotinic acid to bypass defect in NAD synthesis.
Nicotinamine adenine dinucleotide (NAD) is an essential cofactor in many cellular oxidation/reduction reactions and the maintenance of an optimal intracellular concentration of this nucleotide is of paramount importance. In *Escherichia coli* NAD derives from quinolate (QUIN) which is synthesized from L-aspartate in the so-called *de novo* pathway. QUIN is the result of the concerted action of two enzymes, the L-aspartate oxidase (*nadB)*, and the quinolate synthase (*nadA)*. The QUIN is converted by quinolinate phosforibosyl transferase (*nadC*) to nicotinic acid mononucleotide, which enters the pathway for NAD synthesis. In the absence of functional *nadA* and/or *nadB*, exogenous nicotinic acid can be used instead of QUIN to produce nicotinic acid mononucleotide. Loss of the function of any of the three genes *nadB, nadA* can render NAD auxotrophy.
Requirement of nicotinic acid in enteroinvasive *Escherichia coli*. Enteroinvasive *Escherichia coli* (EIEC) represent a midpoint in the evolution between *E. coli* and *Shigella* species. The infection of EIEC is similar to that of *Shigella*, with an intracellular stage of replication and cell-to-cell spread. Based on phenotypic and biochemical evidence EIEC are more similar to *E. coli* than *Shigella*. Moreover EIEC are not a homogenous group respect to serotype, plasmid content, and biochemical features. Recent studies of the genetic relationships between pathogenic and commensal *E. coli* strains, by multilocus enzyme electrophoresis, ribosomal DNA restriction fragment length polymorphism or genome sequencing projects confirm the presence of EIEC among different clusters of *E. coli* species.
Even though *E. coli* can be grouped in the same genus with *Shigella* because of genetic similarities, nicotinamide auxotrophy is observed less frequently in *E.coli* than in *Shigella*. Unlike *Shigella*, the EIEC strains serotype O135 isolated during a study on diarrheal diseases in children in Somalia, were able to grow as prototroph in minimal medium. To verify the nicotinic acid requirement in other EIEC we analyzed several strains isolated in different geographic areas, belong to different serotypes, and display different plasmid contents, but they are all positive in invasivity assay. Among the 15 EIEC strains studied, only four required nicotinic acid supplementation for grown in minimal medium. In these strains the nicotinic acid auxotrophy is due to alterations in *nadB* gene (L-aspartate oxidase), since the introduction of plasmid containing the *nadB* from *E. coli* K-12 restored the nicotinic independence. In addition *nadA* gene of these strains was able to complement *E. coli nadA* mutant. Sequence data of *nadB* from these four strains and site direct mutagenesis showed *nadB* disruption by insertion of an *IS600* element after codon 52 (three strains) or inactivation of L-aspartate oxidase through a change of glycine74 in glutamate (one strain).

Requirement of nicotinic acid in *Shigella*. There are several well known auxotrophic requirements of *Shigella* that are not found among most isolates of *E. coli*. Among these the nicotinic acid supplementation for growth on minimal medium is due to mutation in one or both unlinked loci *nadA* and *nadB,* encoding the enzyme complex that converts L-aspartate to quinolate. Recently as been reported that QUIN is a strong and specific inhibitor of Type three secretion system (TTSS) apparatus. An *in silico* analysis performed on genome sequences currently available on public databases highlights that *nadB* and/or *nadA* are always defective in *Shigella* and that its inactivation has been obtained by diverse strategies. We analyzed a collection of *Shigella* strains isolated over several years in different geographic areas to respect the acid nicotinic requirement. All *Shigella* strains examined required nicotinic acid or the intermediate quinolinic acid suggesting a deficiency in *nadA* or *nad*B locus. Than to analyze the molecular changes responsible of *nadB* or *nadA* silencing we cloned and sequenced the *nadB* or *nadA* loci of *Shigella* (12 *S flexneri* serotype 1-6, one *S dysenteriae 1* and two *S sonnei*) and tested their ability to restore the nicotinic acid prototrophy in *E. coli nadA* or *nadB* mutant. All *Shigella,* except *S. flexneri* serotype 3a and *S. dysenteriae* 1, contained alterations in the *nadB* gene encoding for L-aspartate oxidase, the first enzyme steps in the pathway for *de novo* synthesis of NAD. This is an event that occurs through two strategies: point mutations leading to the replacement of an aminoacid by a stop codon, (*S. flexneri* serotype 1b, 2a, 4a and 5a) or a significant disruption by insertion of *IS600* element (*S. flexneri* serotype 6 and *S. sonnei*). Inactivation of quinolate synthase (NadA) occurs through a change of cysteine 128 in tyrosine (*S. flexneri* serotype 1b, 2a and 4a), and change of prolyne 219 in leucine (*S. dysenteriae* 1) or by insertion of *IS21* element (*S. sonnei*). *S. flexneri* serotype 3a, 5 and 6 contains a functional *nadA* gene able to complement *E. coli nadA* mutant. Since the introduction of a functional *nadB* restore the nicotinic acid independence only in *S. flexneri* serotype 5 the most likely explanation for the absence of quinolate synthase activity in *S. flexneri* serotype 3a and serotype 6 includes the possibility of mutations mapping outside the *nadA* gene that impair the functionality of quinolate synthase in these specific serotype.

All together, our observations confirm and extend the concept that *Shigella* evolution proceeds through convergent evolution toward removal or inactivation of AVL by whatever alteration (IS or point mutation) leads to a more virulent phenotype. Although EIEC strains may be developing the full *Shigella* phenotype, they do not have the full set of characters that define *Shigella* strains. The majority of the EIEC analyzed in this study are nicotinic acid independent and since EIEC retained more characteristics of commensal *E. coli* than *Shigella* spp., these strains might reflect an earlier stage of the evolutionary process undergone by *Shigella*.

RIASSUNTO

Shigella è l' agente eziologico della dissenteria bacillare o shigellosi, che rappresenta un grave problema per la salute pubblica, in particolare nei paesi in via di sviluppo, dove causa più di un milione di decessi l'anno, prevalentemente in bambini di età inferiore ai 5 anni.

Il genere *Shigella* comprende quattro differenti specie: *S. flexneri*, *S. sonnei*, *S. dysenteriae* e *S. boydii*.

Shigella è talmente simile ad *E. coli* (93%omologia) che solo per ragioni storiche viene classificato come genere a parte.

Sono state descritte poche caratteristiche che permettono di distinguere i ceppi di *Shigella* dai ceppi di *E. coli* enteroinvasivi (EIEC), in grado di provocare una sindrome del tutto simile alla dissenteria bacillare. Entrambi i batteri sono caratterizzati dalla presenza di un plasmide di virulenza di grandi dimensioni su cui sono localizzati i geni chiave per l'espressione del fenotipo invasivo. Recenti studi filogenetici hanno suggerito che *Shigella* e EIEC formano un unico patotipo di *E. coli*.

Nella mia tesi di dottorato ho volto la mia attenzione allo studio dell'evoluzione del patotipo *Shigella*/EIEC da un antenato *E. coli* commensale.

Il patotipo *Shigella*/EIEC è evoluto a partire da un ceppo commensale di *E. coli* attraverso due eventi opposti: l'acquisizione di geni di virulenza, localizzati sul plasmide e su isole di patogenicità e la perdita o inattivazione di geni la cui espressione è superflua o interferisce con il meccanismo patogenetico di tali batteri. In *Shigella* ed EIEC sono stati persi geni definiti neutri per la virulenza, ma ridondanti all'interno dell'ospite e geni la cui espressione contrasta la patogenicità e per tale motivo definiti loci di antivirulenza. Di recente sono stati classificati come loci di antivirulenza i geni *nad* necessari per la sintesi *de novo* del cofattore nicotinammide adenina dinucleotide (NAD).

In *E. coli*, due geni *nadA* e *nadB* codificano per il complesso enzimatico che catalizza l'ossidazione di L-aspartato in iminoaspartato, che è poi condensato con il diidrossiacetone fosfato a formare l'acido chinolinico (QUIN), un precursore della sintesi del NAD. In assenza della funzionalità dei geni *nadA* e/o *nadB*, non viene prodotto il QUIN ma il batterio può ancora sintetizzare NAD a partire da sorgenti esogene di acido nicotinico. Una caratteristica biochimica condivisa da tutti i ceppi di *Shigella* è la richiesta di acido nicotinico. Tale richiesta è dovuta a mutazioni nei due geni *nadA* e *nadB*, che codificano rispettivamente per l'enzima chinolato sintetasi e per l'enzima L-aspartato ossidasi. Da recenti studi è emerso che l'acido chinolinico (QUIN) è un potente inibitore della virulenza di *Shigella*.

Analisi genetica dei loci di antivurelenza coinvolti nella sintesi del NAD in ceppi di *Escherichia coli* enteroinvasivi e *Shigella.*

Numerosi studi relativi all'evoluzione di *Shigella* hanno evidenziato che le diverse specie di *Shigella* derivano da *E. coli* in seguito a evoluzione convergente che coinvolge due meccanismi opposti: acquisizione di nuovi geni di virulenza e perdita di geni preesistenti.

Shigella sembra divergere maggiormente da *E. coli* rispetto ai ceppi di *E. coli* enteroinvasivi, un gruppo di *E. coli* che causano dissenteria e che condividono lo stesso meccanismo di patogenicità con *Shigella*. Una caratteristica unica di *Shigella* rispetto ad *E. coli* è la sua stretta dipendenza da una fonte esogena di acido nicotinico dovuta a difetti nel complesso enzimatico coinvolto nella sintesi del NAD.

Il NAD è un cofattore essenziale in tutte le cellule in quanto è necessario in molte reazioni cellulari di ossidoriduzione. Di conseguenza è di fondamentale importanza il mantenimento di una concentrazione intracellulare ottimale di questo nucleotide. La perdita di uno o entrambi i geni che codificano per gli enzimi coinvolti nella sintesi del NAD (*nadB* o *nadA*) rende i ceppi auxotrofi per l'acido nicotinico.

In questo studio, abbiamo analizzato se la dipendenza da acido nicotinico è una caratteristica selezionata nei ceppi EIEC durante l'evoluzione verso lo stile di vita patogeno.

Richiesta di acido nicotinico in *Escherichia coli* enteroinvasivi.

Gli *Escherichia coli* enteroinvasivi (EIEC) rappresentano una fase intermedia nell'evoluzione tra *E. coli* e *Shigella*. L'infezione da EIEC è simile a quella di *Shigella* e consiste nell'invasione delle cellule epiteliali del colon, replicazione intracellulare e diffusione intercellulare. Sulla base di caratteristiche fenotipiche e biochimiche gli EIEC sono più simili a *E. coli* che a *Shigella*. Anche se *E. coli* e *Shigella* possono essere raggruppati nello stesso genere sulla base delle somiglianze genetiche, l' auxotrofia per l'acido nicotinico si osserva meno frequentemente in *E.coli* che in *Shigella*.

Alcuni ceppi EIEC di sierotipo O135 isolati durante uno studio in Somalia, sono in grado di crescere su terreno minimo senza aggiunta di acido nicotinico. Allo scopo di verificare se gli EIEC, diversamente da *Shigella*, siano prototrofi abbiamo analizzato una vasta collezione di ceppi isolati in differenti aree geografiche, appartenenti a differenti sierotipi, con diverso contenuto plasmidico, ma tutti positivi nei saggi di invasività.

Tra i 15 ceppi EIEC analizzati solo 4 richiedevano aggiunta di acido nicotinico per crescere su terreno minimo. In questi quattro ceppi l'auxotrofia è dovuta a mutazioni nel gene *nadB* (L-aspartato ossidasi) in quanto l'introduzione di un plasmide ricombinante contenente il gene *nadB* di *E. coli* K12 ripristina l'indipendenza da acido nicotinic. Inoltre il gene *nadA* di questi ceppi è in grado di complementare un mutante *nadA* di *E. coli.*

L'analisi della sequenza del gene *nadB* di questi quattro ceppi ed esperimenti di mutagenesi sito specifica hanno evidenziato che in tre ceppi il gene *nadB* è interrotto in seguito all'inserzione di un elemento IS*600* dopo il codone 52, mentre in un ceppo l'inattivazione dell'enzima L-aspartato ossidasi è dovuta alla sostituzione della glicina in posizione 74 con acido glutammico.

Richiesta di acido nicotinico in Shigella.

Shigella presenta numerose auxotrofie non presenti nella maggior parte dei ceppi di *E. coli*. In particolare *Shigella* richiede acido nicotinico a causa di mutazioni in uno o in entrambi i geni *nadA* and *nadB* codificano per il complesso enzimatico che converte L-aspartato in acido chinolinico (QUIN). Recentemente è stato dimostrato che il QUIN è un potente inibitore specifico del sistema di secrezione di tipo tre in *Shigella*.

Un'analisi in silico effettuata su sequenze genomiche attualmente disponibili sui database pubblici evidenzia che *nadB* e/o *nadA* sono sempre difettosi in *Shigella* e che la loro inattivazione avviene attraverso diverse strategie.

Abbiamo incluso in questo studio l'analisi di differenti ceppi di *Shigella* ed in particolare 12 *S. flexneri*, 1 *S. sonnei* and 1 *S. dysenteriae* appartenenti a diversi sierotipi e di diversa provenienza geografica allo scopo di individuare i riarrangiamenti molecolari nei geni *nadA* e *nadB* responsabili della richiesta di acido nicotinico in questi batteri.

Tutti i ceppi di *Shigella* esaminati richiedono acido nicotinico o l'intermedio acido chinolinico per crescere in terreno minimo, suggerendo un difetto nei primi passaggi del processo biosintetico che coinvolgono i due loci *nadA* e *nadB*. Allo scopo di analizzare i cambiamenti molecolari responsabili del silenziamento di *nadB* o *nadA*, abbiamo clonato e sequenziato i due loci dai diversi ceppi di *Shigella* e saggiato la loro capacità di ripristinare la prototrofia in mutanti *nadA* o *nadB* di *E. coli*. Tutti i ceppi di *Shigella*, eccetto *S. flexneri* 3a e *S. dysenteriae* 1, mostravano gravi alterazioni nel gene *nadB* che codifica per L-aspartato ossidasi, coinvolto nel primo passaggio della sintesi *de novo* del NAD. Il silenziamento di tale gene si attua attraverso due strategie: mutazioni puntiformi che portano alla sostituzione di un amminoacido con un codone di stop (*S. flexneri* sierotipo 1b, 2a, 4a e 5a) o l'inserzione di una IS*600* a interrompere la sequenza codificante (*S. flexneri* sierotipo 6 e *S. sonnei*).

L'inattivazione dell'enzima chinolato sintasi (NadA) avviene attraverso la sostituzione aminoacidica di un residuo di cisteina in posizione 128 con una tirosina (*S. flexneri* sierotipo 1b, 2a e 4a), il cambiamento di una prolyne in posizione 219 in leucina (*S. dysenteriae 1*) o all'inserzione di una *IS21* nella sequenza codificante (*S. sonnei*).

S. flexneri sierotipo 3a, 5 e 6 contengono un gene *nadA* funzionale, in grado di complementare il mutante *nadA* di *E. coli*. Dal momento che l'introduzione di un gene *nadB* funzionale ripristina l'indipendenza da acido nicotinico solo in *S. flexneri* sierotipo 5, la spiegazione più probabile per l'assenza di attività della chinolato sintasi in *S. flexneri* sierotipo 3a e sierotipo 6 potrebbe essere la presenza di mutazioni che mappano al di fuori del gene *nadA* e che interferiscono con la funzionalità dell'enzima solo in questi specifici sierotipi.

Nel complesso, le nostre osservazioni confermano ed estendono il concetto che l'evoluzione di *Shigella* a partire da *E. coli* commensali procede attraverso la delezione o l'inattivazione di geni di antivirulenza, tramite diverse strategie (IS, delezione o mutazione puntiforme) portando ad un fenotipo più virulento.

Al contrario, poiché i ceppi EIEC sembrano aver mantenuto molte caratteristiche presenti nei ceppi *E. coli* commensali rispetto a *Shigella*, come ad esempio l'indipendenza da acido nicotinico, questi ceppi potrebbero riflettere una fase intermedia nel processo evolutivo che porta da *E. coli* a *Shigella*.

1. INTRODUCTION

Shigella are facultative intracellular pathogens that show a high specificity for human or primate hosts. The first report on the isolation and characterization of bacteria causing bacillary dysentery, later named *Shigella*, was published by Kiyoshi Shiga at the end of the 19th century. *Shigella* causing bacillary dysentery (shigellosis), a severe enteric syndrome, in humans. Shigellosis is endemic throughout the world where it is held responsible for some 150 million cases of severe dysentery with blood and mucus in the stools, the overwhelming majority of which occur in developing countries and involve children less than five years of age (Niyogi, 2005). About 1.1 million people were estimated to die for Shigellosis infection each year, with 60% of the deaths occurring in children under 5 years of age. More recent estimates fix the *Shigella* disease burden at 90 million episodes and 108 000 deaths per year (WHO). Shigellosis is transmitted by the fecal-oral route and is characterized by an acute intestinal inflammation resulting in strong abdominal cramps, fever and bloody diarrhea (Sansonetti, 2006). *Shigella* is traditionally subdivided into four species: *Shigella flexneri*, the prevalent one, causing large endemic infections; *Shigella dysenteriae* responsible for deadly epidemic outbreaks; *Shigella sonnei*, associated with mild infections in industrial countries; and *Shigella boydii*, endemic in the Indian subcontinent (Schroeder and Hilbi, 2008; Yang *et al*., 2005). *Shigella* strains are further divided into 38 serotypes based on O antigen variation: 13 in *S. dysenteriae*, 18 in *S. boydii*, 6 in *S. flexneri*, and 1 in *S. sonnei*. It appears from molecular data that the *Shigella* phenotype arose

within *E. coli*. (Lan and Reeves, 2002).

However, banking on molecular approaches, *Shigella* spp. have been shown to belong to extremely diverse *E. coli* species rather than forming a distinct genus (Pupo *et al*., 2000). The analysis of the evolutionary relationship among *Shigella* and *E. coli* indicates that *Shigella* strains have derived repeatedly from different branches of the *E. coli* tree by convergent evolution involving both gain and loss of genes (Lan and Reeves, 2002; Pupo *et al*., 2000). *Shigella* appears to have diverged from commensal *E. coli* more than enteroinvasive *E. coli* (EIEC), a group of diarrheagenic *E. coli* which shares with *Shigella* the same pathogenicity process. There are very few characteristics that can distinguish *Shigella* strains from enteroinvasive *E. coli* (EIEC), and in the light of this similarity, *Shigella* and EIEC have been grouped into a single pathotype (Kaper *et al*., 2004).

Historically, EIEC was first described in 1944, when it was called paracolon bacillus, but it was later identified as *E. coli* O124. In the 1950s, another group of *E. coli* strains was found to cause experimental keratoconjunctivitis

in guinea pigs by the Sereny test, a trait common with *Shigella*. These strains were placed in the *E. coli* subgroup EIEC as *E. coli* O164 (Rowe *et al.*,1977).

EIEC and *Shigella* spp. bear remarkable phenotypic likeness, with a reduction in the number of substrates utilized relative to commensal *E. coli* strains. These similar phenotypes may be attributed to the fact that these organisms spend much of their lifetime within eukaryotic cells and have a different nutrient supply from most *E. coli* strains (Lan and Reeves, 2002). As compared to commensal and other pathogenic *E. coli* strains, characteristic features of *Shigella* spp. and EIEC strains are the presence of a virulence plasmid (pINV, Figure 1) of ≈220 kb and their ability to induce their entry into epithelial cells and disseminate from cell to cell.

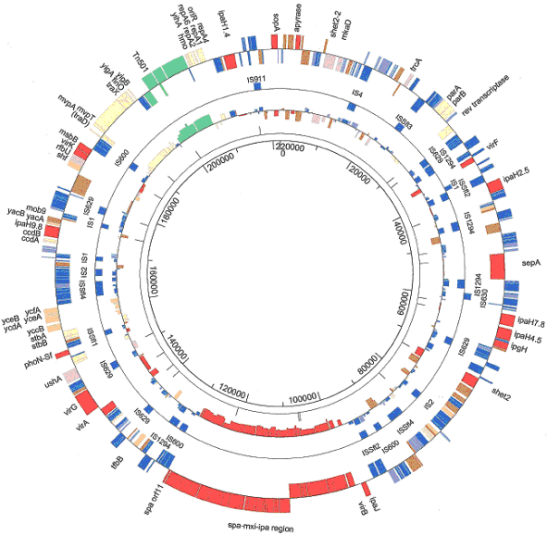

Figure 1. pINV of *Shigella flexneri* serotype 2a (Venkatesan *et al*., 2001).

The virulence plasmid is an essential virulence determinant of *Shigella*/EIEC pathotype and encodes the molecular machinery necessary for tissue invasion and the intracellular lifestyle. The central element of this machinery is a Type Three Secretion System (TTSS). The TTSS enables the bacteria to translocate a set of approximately 25 proteins from the bacterial cytoplasm directly into the eukaryotic host cell, where these "effector" proteins interfere with various host cell processes

1.1 PATHOGENESIS OF *SHIGELLA*

Shigella spp. are transmitted by the fecal-oral route and enter the human body via the ingestion of contaminated food or water. The bacteria are highly infectious, since as few as 10 to 100 microorganisms are sufficient to cause disease (Schroeder and Hilbi, 2008). After passage through the stomach and small intestine, the bacteria reach the large intestine, where they establish an infection.

The pathogenicity of *Shigella* strains is based on their capacity to reach and invade colonic Epithelial Cells (EC), leading to intracellular bacterial multiplication and spread to adjacent cells with consequent cell death and destruction of the colonic mucosa (Sansonetti, 2006; Schroeder and Hilbi, 2008) (Figure 2). In particular, once ingested, these bacteria reach the colon and gain access to the intestinal mucosa by promoting their uptake into M cells in Peyer's patches. The bacteria are then released into an intraepithelial pocket and invade resident macrophages, where they multiply and induce rapid cell death. Released from dead macrophages, enter the surrounding enterocytes through the basolateral surface. Bacterial multiplication within macrophages and enterocytes results in the production of proinflammatory chemokines and cytokines, which in turn leads to polymorphonucleated leukocyte (PMN) infiltration of the intestinal tissue. Migrating PMNs exacerbate the infection since, as they move towards the lumen, they disrupt epithelial tight junctions permitting the bacteria to further access the basolateral surface of enterocytes. Contact with the host cells stimulates the bacterial type three secretion system to deliver several effectors, which facilitate escaping from the host innate defense systems, activate host signalling pathways and induce focused reorganization of the cytoskeletal actin around the bacterial cell.

Inside the host cell, *Shigella* induces lysis of the phagocytic vacuole and the bacterial surface protein IcsA starts to assemble actin tails that propel the bacteria through the cytoplasm and into adjacent cells (Bernardini *et al.*,1989). Proinflammatory signaling by macrophages and EC further activates the innate immune response involving NK cells and attracts PMN. The influx of PMN disintegrates the EC lining, which initially exacerbates the infection and tissue destruction by facilitating the invasion of more bacteria. Ultimately, PMN phagocytose and kill *Shigella*, thus contributing to the resolution of the infection.

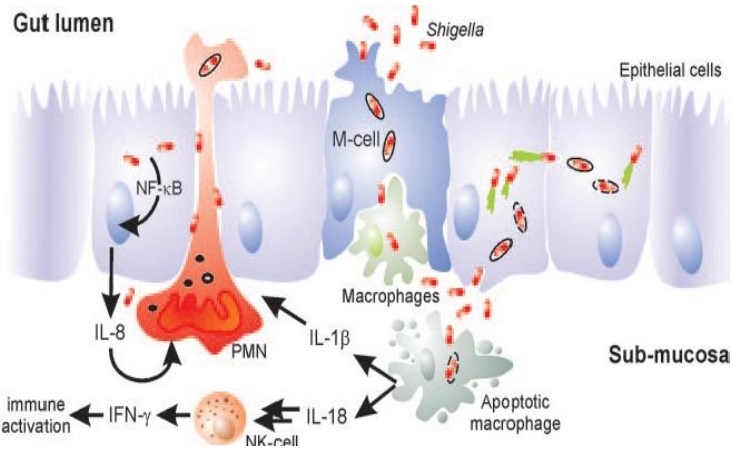

Figure 2. Cellular pathogenesis of *Shigella* spp. (Schroeder and Hilbi, 2008)

The cellular pathogenesis and clinical presentation of shigellosis are the sum of the complex action of a large number of bacterial virulence factors. The essential part of the molecular machinery required for bacterial invasion and intracellular survival is encoded on the large *S. flexneri* virulence plasmid (Sansonetti. *et al.*, 1982; Sansonetti 2001). Sequencing of virulence plasmids from different *Shigella* strains revealed that these plasmids of approximately 200 kb contain a mosaic of around 100 genes and a comparable number of Insertion Sequence (IS) (Buchrieser, C. *et al.*, 2000, Jiang, Y., *et al.*, 2005). The core of the plasmid is the conserved 31-kb entry region, which was shown to be necessary and sufficient for EC invasion and macrophage killing (Sasakawa *et al.*, 1988). The PAI-like region consists of 34 genes that are organized into two clusters transcribed in opposite directions (Figure 3).

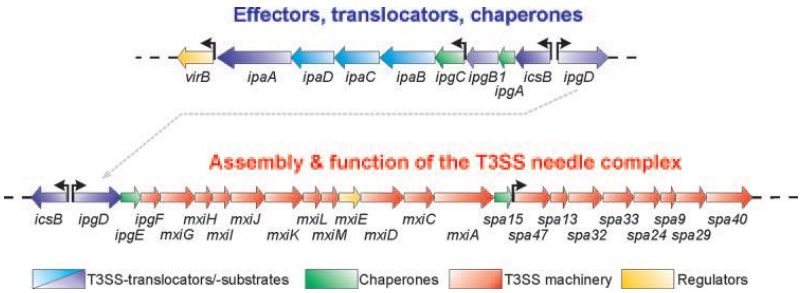

Figure 3. Map of the 31-kb "entry region" on the *S. flexneri* virulence plasmid WR100. The genes indicated encode structural components of the Mxi-Spa TTSS, secreted translocator and effector proteins, chaperones, and regulatory proteins. The region shown is essential and sufficient to invade EC and induce macrophage cell death (Schroeder and. Hilbi, 2008).

Based on their functions, these genes can be divided into four different groups. The first group consists of proteins secreted by the *S. flexneri* TTSS that act as effectors manipulating host cell processes in favor of the bacteria. Among these proteins are the dominant immunogenic antigens of *S. flexneri*, the invasion plasmid antigens IpaA to IpaD (Buysse *et al.*,1987). Three of them, IpaB to IpaD, are the key virulence factors of *S. flexneri*. Apart from having effector functions that are essential for host cell invasion and intracellular survival, these proteins also control the secretion and translocation of other effector proteins into eukaryotic host cells (Menard *et al.*, 1993). Genes of the second group comprise more than half of the entry region and are required for the secretion of the Ipa proteins and other effectors proteins. These genes were designated membrane expression of *ipa* (*mxi*) and surface presentation of *ipa* (*spa*) antigens (Hromockyj and Maurelli. 1989; Venkatesan *et al*., 1992). The *mxi-spa* locus encodes the components needed for the assembly and function of the TTSS, which, together with IpaB, IpaC, and IpaD, allows the direct translocation of effector proteins from the bacterial cytoplasm into the host cell (Blocker *et al.*,1999). Approximately 25 proteins encoded in different locations on the virulence plasmid are secreted via this system.

In addition to the fundamental bacterial weaponry, the entry region contains the two transcriptional activators VirB and MxiE, representing group 3, which regulate TTSS-associated genes located in the entry region or scattered throughout the remainder of the virulence plasmid (Kane *et al.*,2002).

All virulence genes are under the tight control of a regulatory network, which responds to environmental changes encountered upon the entry of the

bacteria into the host. The major virulence plasmid activator, VirF is induced by environmental stimuli and triggers the expression of the central transcriptional activator VirB. VirB promotes the expression of the entry region genes and some additional effector genes scattered on the virulence plasmid. The secretion of the "first-set" TTSS effectors stored in the bacterial cytoplasm enables MxiE/IpgC-controlled induction of a discrete set of already-induced effector proteins and expression of the "second-set" effectors (Figure 4).

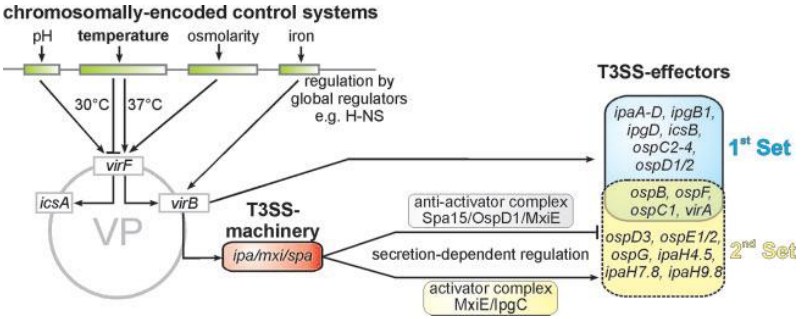

Figure 4. Regulatory elements controlling the expression of the T3SS and its substrates on the *S. flexneri* virulence plasmid. (Schroeder and Hilbi, 2008).

Finally, four genes, classified as group 4, encode chaperones (ipgA, ipgC, ipgE, and spa15) and are also located within the entry region. These chaperones stabilize TTSS substrates in the bacterial cytoplasm, and at least two of them, IpgC and Spa15 participate in the transcriptional regulation of TTSS effector genes located outside of the entry region (Page *et al.*,1999).

1.2 EVOLUTION OF *SHIGELLA*

The past decade has witnessed an explosion of knowledge of the genetic nature of bacterial pathogens thanks largely to the application of whole genome sequencing. The genome of every major bacterial pathogen of human and veterinary importance has now been sequenced. In addition, the genomes of many nonpathogenic bacteria (environmental saprophytes and commensals of human) have been sequenced. These data provide an unprecedented opportunity to examine how bacterial pathogens evolved from their commensal ancestors. A well-established theme in the evolution of bacterial pathogens is acquisition of novel gene traits through horizontal gene transfer (HGT). Many more examples of plasmids that encode virulence determinants that allow pathogenic bacteria to colonize new niches and cause damage to the host have been described (Maurelli; 2007) Bacteriophage genomes, too, are sources of genes that encode traits that allow the newly evolved pathogen to colonize and compete successfully within a new niche (Brussow *et al.*,2004). Of considerable impact to the evolution of bacterial pathogens is the contribution made by pathogenicity islands (PAI). A PAI is a large block of genes (10–200 kb) present in a bacterial pathogen that is missing from a related but nonpathogenic reference strain (Hacker and Kaper, 2000; Schmidt and Hensel, 2004). The island contains genes that are known or suspected to play a role in the virulence of the pathogen. Furthermore, the codon usage of the genes in the island and the atypical G+C content of the DNA in the island relative to the reference genome suggest their foreign origin and acquisition by HGT. Often, one or more PAIs can be identified within the genome of a bacterial pathogen, suggesting that a bacterium can evolve into a pathogen by quantum leaps, i.e. acquisition by HGT of a cluster of virulence genes in the form of a PAI in a single step (Groisman and Ochman, 1996).

Although pathogen evolution progresses through mutation and gene acquisition via HGT, the newly acquired virulence traits, along with the pre-existing core genes, continue to undergo selection after the pathogen has acquired the ability to colonize a new niche. This process is known as pathoadaptation. Pathoadaptive mutations are genetic modifications that enhance the fitness of the pathogen in the novel (host-associated) environment. Pathogenic bacteria also evolve from commensal bacteria by 'loss of function'. Thus, pathoadaptive mutation via gene loss complements the pathway of bacterial pathogen evolution by 'gain of function' mutation and gene acquisition. This model of evolution begins with the premise that genes required for fitness in one niche may actually inhibit fitness in another environment (Maurelli, 2007). Two hypotheses could account for the origin of the different *Shigella*/EIEC groups. These groups could derive from different groups of *E. coli* that acquired the pINV plasmid independently,

either from an unknown donor or from other *Shigella* strains that already had the pINV. Phylogenetic analysis of chromosomal genes, which showed the diversity of genotypes within the *Shigella* genus, was interpreted as evidence for the independent acquisition of the pINV by different *E. coli* groups (Pupo *et al.,* 2000). Alternatively, the pINV plasmid could have been transferred, or constructed, only once in an ancestral *E. coli* strain that subsequently gave rise to the different *Shigella* groups (Escobar Pàramo *et al.,*2003).

Several studies using different technical approaches clearly prove that *Shigella* spp. belong to the species *E. coli*, rather than forming a separate genus (Pupo *et al.,*1997). The sequence divergence between *S. flexneri* and *E. coli* K-12 is about 1.5% (Lan and Reeves. 2002). This is marginal compared to, e.g., 15% in the case of pathogenic *Salmonella enterica*, which is also closely related to *E. coli*. Moreover, diarrheagenic enteroinvasive *E. coli* (EIEC) strains share biochemical characteristics, essential virulence factors, and clinical symptoms with *Shigella* spp. While EIEC does not completely fulfill the definition of the genus *Shigella*, genome analysis revealed a closer relationship to *Shigella* spp. than to commensal *E. coli* strains (Lan and Reeves, 2002).

Comparative genomics clearly indicates that *Shigella* spp. and EIEC evolved from multiple *E. coli* strains by convergent evolution.

Yang and collaborators (Yang *et al.,* 2007) construct a phylogenetic tree using 23 selected chromosomal genes (Figure 5) showing that most of the *Shigella* strains have fallen into three major clusters (C1, C2, and C3), and *S. sonnei* (SS), *S. dysentaeriae* serotype 1, 8, 10, 13 (D1, D8, D10, D13), and *S. boydii* serotype 13 (B13) stand as outliers. SS and D8 are closer than D1, D10, and B13 to the main clusters. The results are consistent with a previous report based on analysis of 8 housekeeping gene sequences (Pupo *et al*. 2000) with increased resolution, indicating that C1 can be subclustered to SC1, SC2, SC3, and a minor branch consisting of only *S. dysentaeriae* serotype 7 (D7). SC1 contains only *S. dysenteriae* strains (D3, D4, D6, D9, D11, D12, and D13), SC2 mainly contains *S. boydii* strains (B1, B3, B6, B8, B10, and B18) but also *S. dysentaeriae* serotype 5 (D5), and SC3 contains three *S. boydii* strains (B2, B4, and B14) and *S. flexneri* serotype 6 (F6). The three EIEC strains do not fall into any *Shigella* cluster but form two independent branches.

Each of the five clusters/subclusters (SC1, SC2, SC3, C2, and C3) contains strains mainly from only one serogroup (Figure 1), which indicates that serological classification is highly correlated with the genotypes and continues to be useful in epidemiologic and diagnostic investigations.

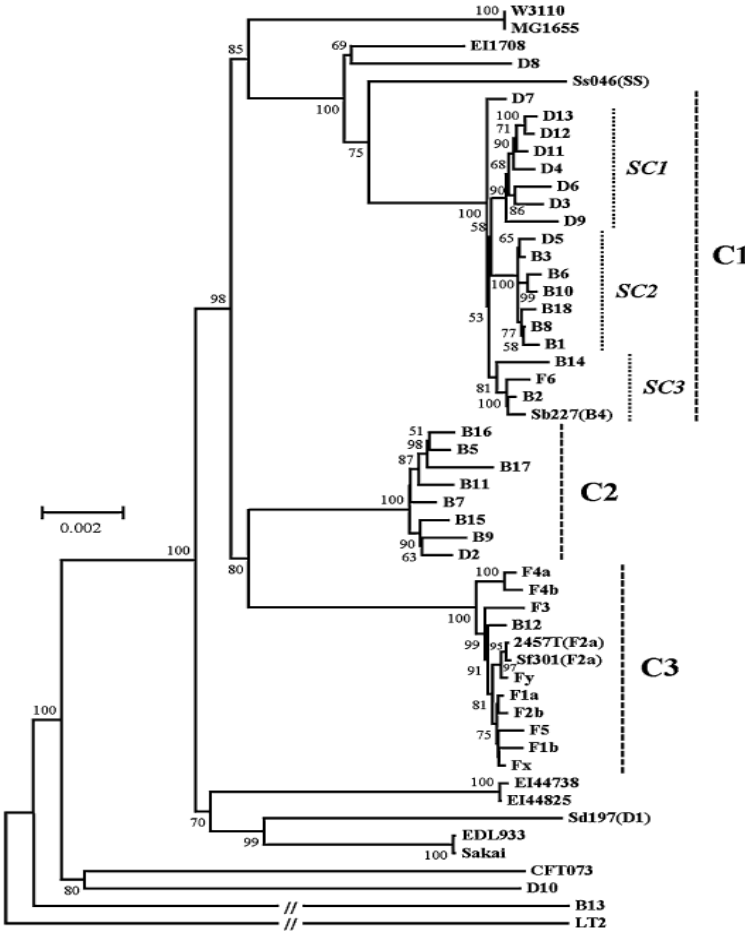

Fig. 5. Phylogenetic tree generated by the neighbor-joining method for the combined data of 23 chromosomal genes. Strain names prefixed by EI are EIEC strains, followed by strain number; for *Shigella* strains, *S. dysenteriae, S. flexneri, S. boydii,* and *S. sonnei* are abbreviated D, F, B, and SS, respectively, followed by the serotype number. Names of those sequenced *E. coli/Shigella* strains were given and their serotypes are indicated in parentheses. *Salmonella* enterica LT2 is used as outgroup. The three major clusters of *Shigella* are indicated by vertical dashed lines at the right, and the vertical dotted lines delineate the newly identified three subclusters of C1 (Yang *et al.,* 2007)

In the latter, three main *Shigella* clusters, each containing strains from the traditionally defined species, are identified. *S. sonnei* and some *S. dysenteriae* strains are more distantly related to these main clusters but still group with *E. coli*. The three main *Shigella* clusters started to diverge from *E. coli* 35,000 to 270,000 years ago (Lan and Reeves, 2002). *S. sonnei* is of more recent origin and separated from the other strains about 10,000 years ago (Shepherd *et al.*,2000). Since EIEC retained more characteristics of commensal *E. coli* than *Shigella* spp., these strains apparently acquired the virulence machinery more recently and might reflect an earlier stage of the evolutionary process undergone by *Shigella* spp. (Lan *et al.*,2004, Yang *et al.*,2005). In addition to a reassignment of the phylogenetic relationships between *Shigella* strains, comparative genomics provides insight into the genetic basis of *Shigella* virulence.

The critical event in the transition toward a pathogenic life-style has probably relied on the acquisition, likely through horizontal transfer, of the virulence plasmid (pINV), which encodes all the genes required for invasion and for intra- and intercellular spreading, including their positive activators. Together with the acquisition of the pINV, the *Shigella*/EIEC pathotype had lost some traits that were important for survival in the environment, but redundant for the life inside the host. The loss of these specific functions has been hypothesized to be functional to the lifestyle of the *Shigella*/EIEC because the expression of these genes, defined antivirulence loci (AVL), might have been detrimental for the expression of the newly acquired virulence genes or because might have been redundant in the new niches (Figure 6)

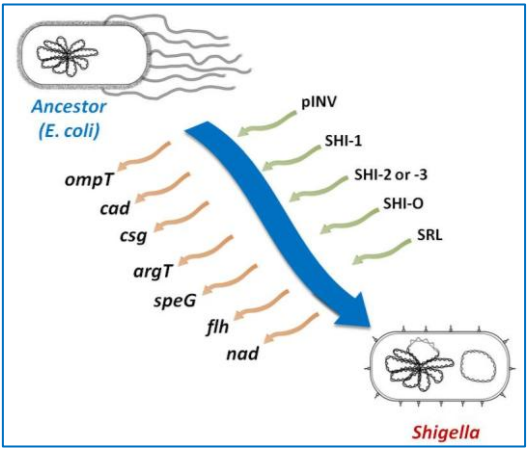

Figure 6. Major steps in the evolution of *Shigella* from ancestral commensal *E. coli*. (Prosseda *et al.*, 2012).

+

The transfer and chromosomal incorporation of large mobile genetic entities carrying one or more virulence-associated genes, termed pathogenicity islands (PAI), are important steps for the rapid evolution of pathogenic bacteria from nonpathogenic progenitors (Dobrindt *et al.*,2004). Both the *Shigella* chromosome and the virulence plasmid reflect these dynamics by containing numerous insertion sequences (IS) and markers of genomic rearrangements (Jin *et al.*, 2002). The distribution of IS types and the phylogenetic grouping of chromosome- and plasmid-encoded genes indicate that the *Shigella* chromosome and the virulence plasmid coevolved (Escobar-Paramo *et al.*, 2003; Lan *et al.*, 2001; Yang *et al.*, 2007). Moreover, these findings suggest that the acquisition and transfer of plasmids between the strains were early steps in the diversification from *E. coli* ancestors. In addition to pathogenicity islands on the virulence plasmid, *Shigella* pathogenicity islands (SHI) were identified in the chromosome (Schmidt and Hensel, 2004). The presence and genomic localization of SHI differ between *Shigella* strains and may contribute to the variation of virulence phenotypes. A specific virulence function has been determined for only some of these genes. The immunoglobulin A-like cytotoxic protease SigA and the enterotoxin ShET1, encoded in SHI-1, were found to induce intestinal fluid accumulation in the rabbit ileal loop model of shigellosis (Fasano *et al.*,1997), and a role for the serine protease Pic in mucus degradation and tissue invasion was postulated based on in vitro studies (Henderson *et al.*, 1999). Moreover,

the SHI-2-encoded aerobactin iron acquisition system was found to contribute to virulence in vivo, and shiA attenuates *Shigella* induced inflammation by suppression of T-cell signaling (Ingersoll *et al.*, 2002, Ingersoll *et al.*, 2003).

A late but important event in the diversification of *Shigella* strains was the acquisition of SHI-O. This PAI harbors genes modifying the structure of the bacterial lipopolysaccharide (LPS) O antigens, thus accounting for the large variety of *Shigella* serotypes (Nie *et al.*, 2006). LPS is a crucial virulence factor, and the host immune response to *Shigella* is serotype specific (Lindberg *et al.*, 1991). The most recent acquisition of genes with new functions is reflected by the emergence of multidrug-resistant *Shigella* strains. A PAI, designated the *Shigella* resistance locus, confers resistance to streptomycin, ampicillin, chloramphenicol, and tetracycline (Turner *et al.*, 2001; Turner *et al.*, 2003).

Besides the extensive gain of virulence factors, the evolution of *Shigella* was shaped by a substantial loss of gene function due to deletion or gene inactivation (Figure 6). Compared to *E. coli* K-12, an average of 726 genes are missing, and over 200 pseudogenes are found per *Shigella* strain (Peng *et al.*, 2006; Yang *et al.*, 2005). Genome reduction due to the loss of functional genes is characteristic of the adaptation of bacteria to an intracellular pathogenic lifestyle (Maurelli, 2007; Moran, 2002). Different factors drive this reductive evolution. The arrival of new virulence factors does not enhance bacterial virulence *per se*, as their expression and function are modulated by the existing genomic background. To be beneficial and therefore retained, virulence factors might require a specific genomic background (Escobar-Paramo *et al.*, 2003). Consequently, genes that attenuate or interfere with virulence are prone to be inactivated. *Shigella* prominently exemplifies the inactivation of so-called "antivirulence genes" (Figure 6).

It is well known that *Shigella* spp. are not motile, as they have lost the ability to synthesize flagella.

The lack of motility is a result of convergent evolution, mainly due to IS mediated inactivation of the flagellar master operon flhDC.

Clear evidence for a direct correlation between lack of motility and virulence in *Shigella* does not exist. Considering that, by HGT, *Shigella* has acquired an efficient intracellular actin-based motility system (Bernardini *et al*., 1989), flagellar motility might have become superfluous and, via a reductive evolution process, flagellar operons might successively have been pruned from the genome. Another surface structure which, as opposed to *E. coli*, is not found in *Shigella* is represented by the so-called curli, a peculiar class of thin aggregative fimbriae encoded by csg genes (Sakellaris *et al*., 2000). Due to their ability to mediate formation of biofilms and to bind to many extracellular proteins in commensal and pathogenic *E. coli*, curli have

been considered effective virulence factors (Barnhart and Chapman, 2006). Despite this role, in *Shigella* the csg operons have been silenced. Multiple IS elements have often interrupted csg genes, triggering genetic rearrangements which have led to small and large deletions. Since, in *Shigella* and EIEC, strong selective pressure against the maintenance of curli exists, it has been hypothesized that the lack of curli might constitute a further case of pathoadaptive mutations (Sakellaris *et al*., 2000). Nevertheless, as for flagella, consolidated experimental proof in favor of a relationship between loss of curli and increased virulence is not available. On the other hand, it is important to recall that surface structures like flagella and curli are well known as potent activators of the host immune system and that their loss might have significantly contributed to minimizing exposure to the host defense system.

1.3 PATHOADAPTIVE MUTATIONS

The evolution of bacterial pathogens from their nonpathogenic ancestors is thought to proceed mainly through the acquisition of virulence genes, often located on plasmids, bacteriophages or pathogenic islands (PAIs) (Hacker and Kaper, 2000). The acquired virulent traits allow a bacterium to probe new host environments for the successful establishment of a pathogenic lifestyle. Then, colonization of the host exerts a strong selective pressure on the optimization of virulence features. This implies the emergence of several changes in the ancestral genome which often redesign important metabolic properties, including the loss of catabolic pathways (Ochman and Moran, 2001). While examples of the evolution of microbial pathogens by horizontal gene transfer (HGT)-mediated acquisition of genes are widespread, the contribution of gene loss to the evolution of pathogenicity is just beginning to be taken into consideration.

Many AVGs such as those responsible for lysine decarboxylase activity, for example the surface protease encoded by *ompT* and quinolinic acid synthesis, have been identified in *Shigella* (Figure 7).

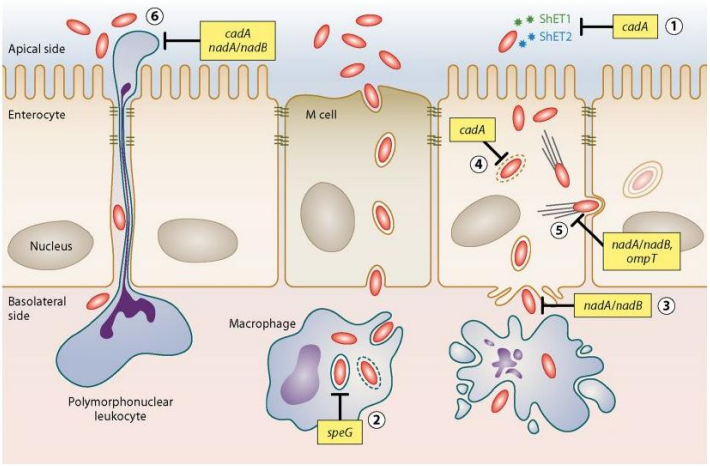

Figure 7. Inhibition of pathogenesis in *Shigella*. Pathogenesis phenotypes interrupted by *Shigella* AVGs (Bliven and Maurelli, 2012)

OmpT. Evidence for the role played by gene loss in the invasive phenotype of *Shigella* initially came from the observation that replacement of the *S. flexneri kcpA* locus with the corresponding region of *E. coli* K-12 hampers the ability of *Shigella* to cause keratoconjunctivitis in guinea pigs (Formal *et al.*,1971). The function of the kcpA locus remained undefined until Nakata *et al.* (1993) discovered that kcpA was not a virulence determinant, but a wide deletion in the *Shigella* genome. At that position, the *E. coli* K-12 chromosome contains remnants of the lambdoid phage DLP12 (Lindsey *et al.*,1989) which carry the ompT gene encoding a surface protease. In contrast to *E. coli* K-12, all *Shigella* spp. and EIEC strains lack the DLP12 prophage and fail to produce the OmpT protease (Nakata *et al.*,1993). Loss of the ompT gene can be considered a pathoadaptive mutation, since the introduction of OmpT negatively interferes with the invasive process by drastically reducing the ability of *Shigella* to spread into adjacent epithelial cells. In particular, in *Shigella*, the OmpT protein triggers degradation of IcsA, a pINVencoded outer membrane protein required for full expression of the invasive phenotype (including keratoconjunctivitis in guinea pigs) and for intracellular bacterial motility (Bernardini *et al.*,1989). It is as yet unclear whether *E. coli* lineages that gave rise to *Shigella* did not harbor DLP12 ab initio or whether the prophage was excised during the pathoadaption process.

Cad. Another well known example of pathoadaptive mutation in *Shigella* is represented by the cad genes (Maurelli *et al.*,1998). They code for a major amino acid decarboxylase system involved in maintaining the pH at a level suitable for cell survival (Foster, 2004). In *E. coli*, the cad locus encodes lysine decarboxylase (CadA), a lysine-cadaverine antiporter (CadB) and a regulator of the cadBA operon (CadC). The end product of lysine decarboxylation is cadaverine, a small polyamine synthesized under the positive control of CadC in response to acid stress. Due to the lack of lysine decarboxylase (LDC^- phenotype), *Shigella* and EIEC do not produce cadaverine. Despite its important role in protecting cells from acidification, cadaverine negatively interferes with the pathogenicity of *Shigella*. In particular, cadaverine seems to induce the attenuation of enterotoxicity and also to inhibit migration of polymorphonuclear leukocytes (PMN) across the intestinal epithelium (Maurelli *et al.*,1998; McCormick *et al.*,1999).

Moreover, cadaverine retards the lysis of the *Shigella*-containing vacuole, thus preventing the bacterium from adequately interacting with the host cell cytoskeleton (Fernandez *et al.*,2001). This evidence taken together indicates that, in *Shigella*, silencing of the cad locus and consequent elimination of cadaverine production represent an important pathoadaptive mutation necessary for increasing the pathogenic potential of bacteria in host tissues (Maurelli, 2007; Prosseda *et al.*,2007).

Comparative genome analysis reveals that within the *Shigella*/EIEC pathotype, convergent evolution has operated independently several times to prevent synthesis of cadaverine.

While, in *S. flexneri*, remnants of the *cadB* gene confirm the past existence of a *cad* locus, in *S. sonnei* and EIEC strains the cad region continues to maintain colinearity with the *E. coli* K-12 chromosome, but is silenced by the insertion of IS elements (Casalino *et al.*,2003, 2005; Maurelli *et al.*,1998). More drastic rearrangements have been detected in S. boydii, where the cad locus is replaced by a P4-like prophage, or in some strains of *S. flexneri* and EIEC, where the cad locus is completely absent (Casalino *et al.*,2003; Day *et al.*,2001).

Interestingly, the absence of CadC in *Shigella* increases expression of arginine decarboxylase, a key enzyme of the arginine-dependent acid resistance system (Foster, 2004), thus partially counterbalancing the loss of lysine decarboxylase (Casalino *et al.*,2010).

Recent evidence suggests that the antivirulence role of the cad operon is not limited to the *Shigella*/EIEC pathotype, but rather also applies to other pathogenic *E. coli* (Torres, 2008).

ArgT. Another interesting issue which underlines how silencing a specific function promotes full expression of virulence in *Shigella* concerns the *argT* gene, which encodes a lysine/ arginine/ornithine-binding protein localized in

the periplasm (Zhao *et al.*,2010). Proteome analysis reveals that, in *S. flexneri* at 37° C, the ArgT level is dramatically reduced compared to 30° C, while the mRNA level does not appreciably vary. The reduction in ArgT depends upon its degradation by HtrA (known also as DegP), a periplasmic protease. Interestingly, despite the fact that *E. coli* and *S. flexneri* ArgT differ only by three amino acids, the *E. coli* protein is not submitted to HtrA degradation either in an *E. coli* or in a *S. flexneri* background. Tyr225 is the key residue which renders *Shigella* ArgT (ArgTSf) susceptible to proteolytic attack.

Competitive invasion assays carried out in HeLa cells and in BALB/c mice using *S. flexneri* strains either lacking or overexpressing ArgTSf indicate that the presence of ArgTSf attenuates the virulence of *Shigella*. Though the basis of this phenomenon remains largely unknown, it can be argued that a possible target of ArgTSf is HtrA itself, since at 37° C its abundance is reduced upon ArgTsf overexpression (Zhao *et al.*,2010). HtrA contributes to the *Shigella* pathogenicity process, as it is required for efficient polar localization of IcsA, the *Shigella* protein responsible for intra- and intercellular movement (Purdy *et al.*,2007). Both ArgT and HtrA are located in the periplasmic space and it is therefore possible that ArgTsf, if not degraded by HtrA, might negatively interfere with HtrA. Another possibility for explaining interference of ArgTsf upon full expression of *Shigella* virulence is that the amino acids transported by ArgT reduce virulence (Zhao *et al.*,2010). In this context, it is worth mentioning that in *Shigella*, ornithine in conjunction with uracil is able to significantly reduce T3SS expression as well as hemolytic activity (Durand and Bjork, 2009).

It is intriguing to speculate that the temperature-dependent silencing of ArgTsf could be a novel strategy of pathoadaptation: outside the human host, amino acid transport in *Shigella* would rely on functional ArgTSf, while within host cells ArgTSf would be degraded to prevent it from interfering with virulence.

SpeG. A recent entry in the list of pathoadaptive mutations in *Shigella* is the *speG* gene, which encodes spermidine acetyltransferase (SAT) (Barbagallo *et al.*, 2011). SAT catalyzes the conversion of spermidine into a physiologically inert form, acetylspermidine. Analysis of the polyamine content reveals that, compared to *E. coli* K-12, *Shigella* spp. accumulate spermidine intracellularly, whereas acetylspermidine and cadaverine are completely absent. In analogy with the lack of *cad* and *nad* genes, also in the case of speG, loss of functionality is the result of convergent evolution. Comparative analysis of the *speG* locus of several *Shigella* strains reveals that inactivating events range from point mutations to gene deletions. In particular, the entire *speG* locus, with only a missense mutation, is present in some *S. dysenteriae* strains, while it is completely absent in all *S. sonnei* and in some strains of *S. flexneri* and *S. boydii.*

Between these two extremes, a series of genetic rearrangements, mainly caused by IS transposition, have been observed, confirming the existence of strong selective pressure towards the loss of SpeG.

The pathoadaptive nature of speG inactivation has been demonstrated by observing that spermidine accumulation increases resistance to oxidative stress and allows *Shigella* to better face the adverse conditions it encounters within macrophages (Barbagallo *et al.*,2011). Indeed, when a functional speG gene is reintroduced in *Shigella*, acetylspermidine is produced, thus reducing the intracellular pool of spermidine. This, in turn, decreases survival of the SpeG-complemented strains within macrophages and increases their sensitivity to oxidative stress conditions. As one of the first steps in *Shigella* infection, survival within macrophages is crucial for successful invasion (Ray *et al.*,2009). It is therefore not surprising that, during pathoadaptation, *Shigella* has altered its polyamine profile to improve resistance to the hostile conditions of the macrophage environment.

Interestingly, transcriptomic analysis of an *E. coli* K-12 strain carrying the *Shigella* VirF regulator indicates that speG expression is upregulated by VirF, the pINV-encoded primary activator of the invasive regulon.

This suggests that acquisition of the pINV by HGT might have increased expression of *speG* in *Shigella*, thus drastically reducing the spermidine content. This could have favored the spread of *speG* mutants which, due to their higher spermidine content, would have been able to better survive within macrophages.

The drastic changes in polyamine content in *Shigella* spp. as compared to *E. coli* stress the relevance of their intracellular level for full expression of the virulence program.

Nad. In order for *E. coli* to synthesize nicotinic acid mononucleotide, a precursor of the essential coenzyme NAD, the enzymes L-aspartate oxidase (encoded by *nadB*) and quinolinate synthetase A (encoded by *nadA*) are required. In contrast, *Shigella* has a strict nutritional requirement for nicotinic acid for growth on minimal media, suggesting that *Shigella* relies on a salvage pathway for NAD (Gemski P, *et al.*, 1971). When the sequences of genes required for *de novo* NAD synthesis in *E. coli* were compared to the *Shigella* genome, it quickly became apparent that all *Shigella* species have inactivated *nadA*, *nadB*, or both genes (Prunier *et al*. 2007a, Prunier *et al*. 2007b). These inactivation take the form of multiple amino acid substitutions, complete or partial deletions of *nadA* or *nadB*, or insertion sequence (IS) elements interrupting either or both genes. The wide variety of different genetic alterations utilized by *Shigella* to inactivate these genes suggests that there was strong selective pressure to lose this part of the NAD pathway. To compensate, *Shigella* bypasses this early block in the pathway by importing exogenous nicotinic acid and converting it to the NAD precursor nicotinic

acid mononucleotide through the actions of nicotinate phosphoribosyltransferase (Prunier *et al*., 2007a). Prunier *et al*. showed that one of the intermediates in the *de novo* synthesis pathway, quinolinic acid, attenuates both *Shigella* invasion and intracellular dissemination, in addition to blocking PMNtransepithelial migration (Prunier *et al*., 2007a). The secretion of *Shigella* type III secretion system (TTSS) effectors, such as IpaB and IpaC, is greatly decreased in the presence of quinolinic acid, suggesting that inadequate secretion of *Shigella* effectors most likely contributes to these phenotypes. Interestingly, this inhibition is limited to the *Shigella* TTSS. Other organisms with a TTSS, such as *Salmonella enterica*, enteropathogenic *E. coli*, or *Y. enterocolitica*, do not exhibit virulence inhibition in the presence of quinolinic acid. The exact mechanism of quinolinic acid inhibition upon *Shigella* virulence is not yet known.

1.4 THE HARMFUL EFFECTS OF QUINOLINATE (QUIN), AN INTERMEDIATE IN NAD SYNTHESIS

Nicotinamide adenine dinucleotide (NAD) and NAD-phosphate (NADP) are compounds of immeasurable importance in cellular metabolism. They function in numerous anabolic and catabolic reactions and are widely distributed throughout biological systems **(**Foster *et al.*,1980)

NAD and NADP are known to participate in over 300 enzymatically catalyzed oxidation-reduction reactions. In addition, a number of reactions have been discovered in which NAD serves as a substrate. For example, certain procaryotes, such as *Escherichia coli*, utilize NAD as a substrate for deoxyribonucleic acid ligase, an essential for deoxyribonucleic acid synthesis, repair, and recombination. NAD also serves as a substrate in reactions that produce poly-adenosine 5'-diphosphate-ribose. Adenosine 5'-diphosphate ribosylation is proving to be of great importance to both eukaryotic and procaryotic cells. Reduced pyridine nucleotide coenzymes also play an important role in the regulation of amphibolic pathways, such as the citric acid cycle and the oxidative pentose pathway (Foster *et al.*,1980). Extensive research on the biosynthesis of NAD reveal two main biosynthetic pathways to NAD. One pathway involves the aerobic degradation of tryptophan by mammalian cells and a number of lower eucaryotes. Another pathway, found predominantly in procaryotes, is anaerobic and utilizes low-molecular-weight precursors for the synthesis of the pyridine ring structure of NAD (Figure 8). Both pathways lead to the formation of quinolinic acid (QA). Subsequent conversion of QA to NAD occurs via a pathway common to all organisms that have been examined to date **(**Foster *et al.*,1980).

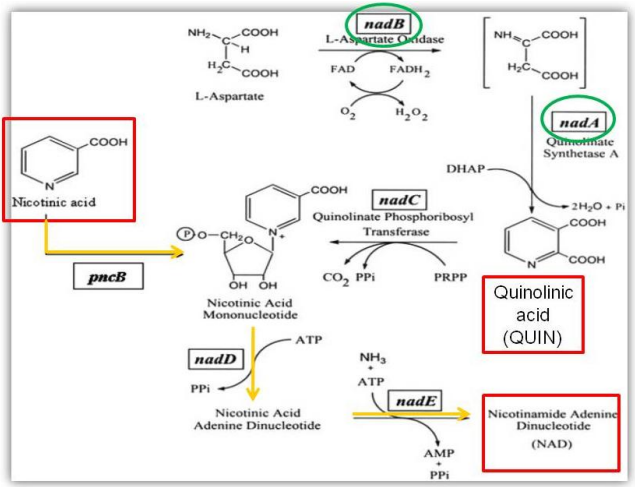

Figure 8: The biosynthetic pathways of nicotinamide adenine dinucleotide (NAD).

The L-aspartate-dihydroxyacetone phosphate pathway is used by *E. coli* K-12 for *de novo* synthesis of NAD (Figure 8). L-Aspartate oxidase, encoded by *nadB*, forms a multi-enzyme complex with quinolinate synthetase A, the product of *nadA*. Quinolinate synthase A catalyzes a unique condensation reaction between dihydroxyacetone phosphate and iminoaspartate, yielding quinolinic acid (QUIN), a central intermediate in the biosynthesis of nicotinamide adenine dinucleotide and its derivatives. Quinolinate phosphoribosyltransferase, the product of *nadC*, converts QUIN to nicotinic acid mononucleotide, which enters the pathway for NAD synthesis. The *nadA*, *nadB* and *nadC* loci have been mapped by transduction and are located at approximately 17, 49 and 1.5 min on the *E. coli* chromosome, respectively (Griffith R. *et al.*,1975).

NadA from *E. coli* contains nine cysteine residues which can be classified as follows: three strictly conserved (Cys113, Cys200 and Cys297) and six partially conserved (Cys64, Cys119, Cys128, Cys195, Cys291 and Cys294) residues. Since NadA contains a CXXCXXC motif, it was dimostrated that the three cysteines of the C113XXC200XXC297 motif were the iron ligands of the [4Fe–4S] cluster and Cys291 and Cys294 form a disulfide bridge and are important for activity (Rousset *et al.;* 2008). Cys64, Cys119, Cys128, Cys195 are not important for the iron ligands, in fact changing this residue with an aminoacid with the same structure, for example serine, didn't affect the functionality of the proteins (Saunders *et al.,* 2008). Biogenesis of iron-

sulfur (Fe-S) cluster-containing proteins relies on assistance of complex machineries. In *E. coli*, has been demonstrated the involvement of the ISC system in the Fe-S cluster formation: ISC system is synthesized under normal growth conditions and is viewed as fulfilling a housekeeping function. Defects in this system, and particularly in *iscS* gene, alter functioning of several Fe-S enzymes, leading to mutated phenotypic traits including, among others, nicotinic acid and vitamin B1 auxotrophies. In *Escherichia coli*, the ISC components are encoded on the iscRSUA polycistronic transcript. Inactivation of *iscS* produces severe growth defects, indicating its biological importance (Takahashi and Nakamura, 1999; Schwartz *et al*, 2000).

The gene *nadB* is present in several microorganisms and in plants, but the protein has been purified only from *Escherichia coli*, Pyrococcus horikoshii and Sulfolobus tokadaii, and characterized from a biochemical and structural point of view only from *E. coli* and *S. tokadaii*. It is a flavoprotein able to use different electron acceptors such as oxygen, fumarate, cytochrome C and quinones suggesting that it is involved in NAD biosynthesis in anaerobic as well as aerobic conditions. This enzyme produces iminoaspartate, which is then converted to quinolinate through the condensation with DHAP catalyzed by NadA.

In the *Escherichia coli* and *Salmonella typhimurium* NAD biosynthetic pathways, the genes involved in the *de novo* synthesis are under negative transcriptional control by the product of the *nadR* locus (Holley B. A. *et al.*, 1985; Tritz, G. J., and J. L. Chandler 1973; Zhu, N. *et al.*, 1988). It has been demonstrated that the *nadR* gene product is a bifunctional protein that can both act as a repressor and control the transport of exogenous nicotinamide mononucleotide (NMN), the immediate NAD precursor, across the cytoplasmic membrane (Foster, J. W. *et al.*,. 1990; Zhu, N., and J. R. Roth. 1991). The latter function would be exerted through a regulatory modulation of the activity of the integral membrane protein PnuC, which is responsible for NMN transport (Foster, J. W. *et al.*,. 1990, Zhu, N. *et al.*,. 1991). While the repression function resides in the NadR N-terminal domain, the C-terminal domain is involved in the transport function (Foster, J. W. and T. Penfound 1993). Both functions appear to exert their control in response to intracellular NAD (or NADP) levels (Foster, J. W. *et al.*,. 1990, 18). In this regard, it has been proposed that NadR behaves as an allosteric protein: in the presence of high NAD levels it might assume a conformation which allows both repression of NAD biosynthetic genes transcription and inhibition of NMN transport system; conversely, when NAD levels are low, NadR might associate with the membrane, allowing full expression of biosynthetic genes and stimulating NMN uptake (Foster, J. W. *et al.*,. 1990, Zhu, N. *et al.*, 1991) (Figure 9).

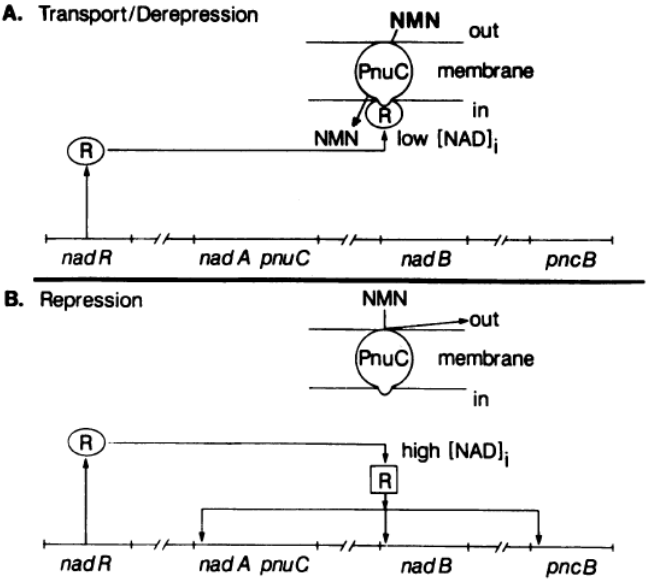

Figure 9. Hypothesized model for the bifunctional nature of NadR. (A) Under low internal NAD conditions (low [NAD],), NadR (R) collaborates with PnuC in the transport of NMN. (B) Under a high internal NAD concentration (high [NAD]i), NadR assumes a conformatio [R] that increases its affinity for nad operator sequences but is less suitable for transport (Foster, J. W. *et al.*,. 1990).

In the absence of functional *nadA* and/or *nadB*, QUIN is not made and no nicotinic acid mononucleotide is available to synthesize NAD. However, exogenous nicotinic acid can be converted to nicotinic acid mononucleotide by the action of nicotinate phosphoribosyltransferase, the product of pncB. This pathway bypasses the need for *nadA* and *nadB* and synthesis of QUIN and is the basis of the nicotinic acid requirement for *Shigella* growth on minimal medium (Gemski *et al.*,1971).

In fact , unlike *E. coli*, *Shigella* spp. lack the *de novo* pathway for synthesis of NAD and therefore require nicotinic acid for growth on minimal medium. To ascertain whether nicotinic acid auxotrophy is the result of a pathoadaptive process, Prunier and collaborators (Prunier *et al.,* 2007a) analyzed the capacity of several intermediates in the NAD biosynthetic pathway to interfere with full expression of virulence. Their observations

show that QUIN is the only intermediate affecting virulence and that it is involved in the inhibition of several steps in the pathogenicity process. In particular, the presence of QUIN decreases intercellular spreading of *Shigella*, either by blocking the ability of bacteria containing protrusions to invade adjacent cells or by inhibiting bacterial ability to lyse protrusion and move to adjacent cells. In turn, this severely reduces transepithelial migration of PMNs. The antivirulent role played by QUIN is further confirmed by its capacity to inhibit secretion of the IpaB and IpaC effectors through the *Shigella* TTSS system (Prunier *et al.,* 2007a).

2. RESULTS

The analysis of the evolutionary relationship among *Shigella* and *E. coli* indicates that *Shigella* strains have been derived repeatedly from different branches of the *E. coli* tree by convergent evolution involving both, gain and loss of genes (Lan and Reeves, 2002; Pupo *et al.*, 2000; Peng *et al.*, 2009). *Shigella* appears to have diverged from commensal *E. coli* more than enteroinvasive *E.coli* (EIEC), a group of diarrhogenic *E. coli* which shares with *Shigella* the same pathogenicity process. A characteristic of *Shigella*, is their strict requirement for an exogenous source of nicotinic acid to bypass defect in NAD synthesis.

Nicotinamine adenine dinucleotide (NAD) is of central importance to cellular metabolism. It functions as an essential cofactor in many cellular oxidation/reduction reaction. Consequently the maintenance of an optimal intracellular concentration of this nucleotide is of paramount importance. In *Escherichia coli* NAD derives from quinolate which is synthetized from L-aspartate and dihydroxyacetone phosphate (DHAP) in the so-called de novo pathway. Quinolate (QUIN) result of the concerted action of two enzymes, the L-aspartate oxidase encoded by *nadB* gene, and the quinolate synthetase encoded by *nadA* gene. The QUIN is converted by quinolinate phosforibosyl transferase (*nadC*) to nicotinic acid mononucleotide, which enters the pathway for NAD synthesis. In the absence of functional *nadA* and/or *nadB*, exogenous nicotinic acid can be used instead of QUIN to produce nicotinic acid mononucleotide (figure 10). Loss of the function of any of the three genes *nadB, nadA* and *nadC* can render NAD auxotrophy.

In this study, we will analyse if the auxotrophic requirement for nicotinic acid may be selected in EIEC strains, which are as regarded as evolutionary intermediates from *E.coli* to *Shigella*.

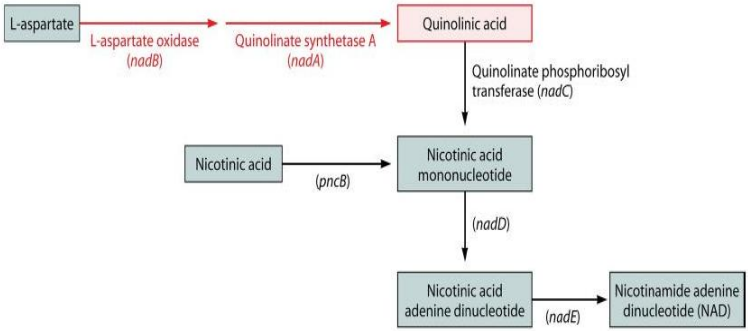

Figure 10: Biosynthetic and salvage NAD pathway in *Shigella*. Compounds or enzymes still present in *Shigella* are marked in black; those that have been lost are marked in red. (Bliven and Maurelli, 2012)

2.1 REQUIREMENT OF NICOTINIC ACID IN ENTEROINVASIVE *E. COLI*

Enteroinvasive *Escherichia coli* (*EIEC*) represent a midpoint in the evolution between *E. coli* and *Shigella* species. The infection of enteroinvasive *E. coli* is similar to that of *Shigella,* with an intracellular stage of replication and cell-to-cell spread. Based on phenotypic and biochemical evidence *EIEC* are more similar to *E. coli* than *Shigella*. Moreover *EIEC* are not a homogenous group since EIEC strains differ with respect to serotype, plasmid content, and biochemical features. Moreover, recent studies of the genetic relationships between pathogenic and commensal *E. coli* strains, by multilocus enzyme electrophoresis, ribosomal DNA restriction fragment length polymorphism or genome sequencing projects confirm the presence of EIEC among different clusters of *E. coli* species (Pupo *et al.,* 2000). Even though *E. coli* can be grouped in the same genus with *Shigella* because of genetic similarities (Lan *et al.*,2004), nicotinamide auxotrophy is observed less frequently in *E.coli* than in *Shigella* (Bergthorsson and Roth, 2005; Prunier *et al*., 2007a) Unlike *Shigella* the EIEC strains serotype O135 isolated during a study on diarrheal diseases in children in Somalia **(**Casalino *et al*., 1994; Nicoletti *et al*., 1988**)** were able to grow as prototroph in M9 minimal salt medium (Table 1).

To verify the nicotinic acid requirement in other EIEC we analyse strains obtained from the Reference Centers of the Pasteur Institute and of the Walter Reed Army Institute of Research. The EIEC strains used in this study (Table 1) were isolated in different geographic areas, belong to different serotypes, and display different plasmid contents, but they are all positive in invasivity assay. First we analyse the ability of EIEC strains to grow in minimal medium. Overall, among the 15 strains studied, only four require nicotinic acid supplementation for grown in minimal medium, namely 4608, 1380, FG4 and FG5

TABLE 1: Bacterial strains used in this study

Strains	**Relevant features**	**MM[a]**	**Country of origin or source (reference)**
MG1655	*E. coli K12; F– λ–ilvG rfb-50 rph-1*	+	Blattner *et al.*, 1997
JW0733	*nadA* deletion mutant of *E. coli* K-12	-	Japan (Baba *et al.*, 2006)
JW2558	*nadB* deletion mutant of *E. coli* K-12	-	Japan (Baba *et al.*, 2006)
HN280, HN11, HN13, HN19, HN300	*EIEC* O135	+	Somalia (Nicoletti *et al.*, 1988).
4608	*EIEC* O143	-	WRAIR (Sansonetti *et al.*, 1982)
53638	*EIEC* O144	+	WRAIR (Sansonetti *et al.*, 1982)
1380	*EIEC* O124	-	IPC
6.81	*EIEC* O115	+	IPC
FG1	*EIEC* NA	+	IPC
FG2	*EIEC* O124	+	IPC
FG4	*EIEC* O164	-	IPC
FG5	*EIEC* O136	-	IPC
FG6	*EIEC* O152	+	IPC
FG7	*EIEC* O144	+	IPC
SFZM50	*S. flexneri* 1b	-	Somalia (Casalino *et al.*, 1994)
SFZM46 YSH6000	*S. flexneri* 2a	-	Somalia (Casalino *et al.*, 1994) Japan (Sasakawa *et al.*, 1986)
SFZM7 SFZM49	*S. flexneri* 3a	-	Somalia (Casalino *et al.*, 1994)
SFZM8, SFZM15 SFZM53	*S. flexneri* 4a	-	Somalia (Casalino *et al.*, 1994)
M90T	*S. flexneri* 5a	-	WRAIR (Sansonetti *et al.*, 1982)
SFZM2,SFZM9 SFZM43	*S. flexneri* 6	-	Somalia (Casalino *et al.*, 1994)
SSZM34 SSZM45	*S. sonnei*	-	Somalia (Casalino *et al.*, 2005)
SDZM603	*S.dysenteriae 1*	-	Somalia (Barbagallo *et al.*, 2011)

MM[a] grow in M9 minimal medium.
NA: not applicable;
WRAIR: Walter Reed Army Institute of Research;
IPC: Institut Pasteur Collection;

In all the four EIEC strains the auxotrophy is satisfied by the intermediate quinolinic acid, suggesting a deficiency in one of the first two steps in the pathway: L-aspartate oxidase (encoded by *nadB*) or quinolate synthase (encoded by *nadA*). In these strains the nicotinic requirement is due to defect in *nadB* gene since the recombinant plasmid containing *nadA* gene of these EIEC strains ($pNadA_{4608,}$ $pNadA_{1380}$,$pNadA_{FG4}$ and pNadA $_{FG5}$) is able to complement JW0733 *E. coli nadA* mutant in minimal medium (Table 2). In addition the introduction of recombinant plasmid containing the wild–type *nadB* gene of *E. coli* MG1655 ($pNadB_{mg}$) restore the nicotinic independence (Table 2).

Than to analyse the molecular rearrangement responsible of *nadB* silencing of the four EIEC strains (4608, 1380, FG4 and FG5) we cloned and sequenced the *nadB* locus of these EIEC strains. Moreover amplification performed using a primer pair spanning the entire *nadB* and genomic DNA from these four EIEC strains as template, showed that only 4608 generates an amplicon with the same size (2.0 kb) of the control strain MG1655. The other three EIEC strains gave products larger than MG1655: 1380 and FG4 give a product of about 3.2 kb, FG5 a product 3.9 kb (Figure 11) probably due to insertion element.

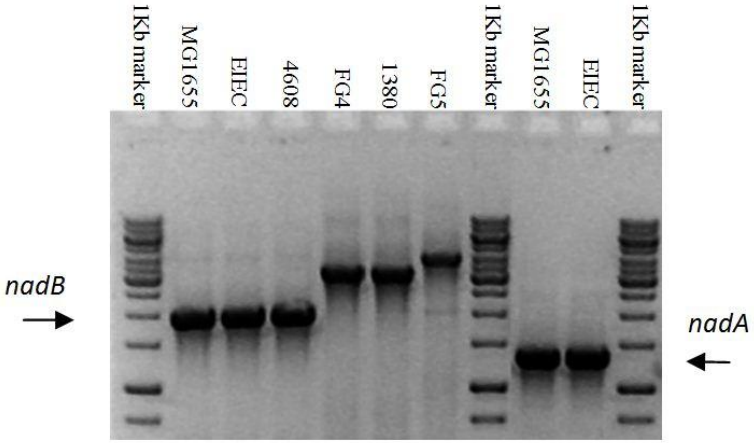

Figure 11: Agarose-gel electrophoresis of the *nadA* and *nadB* PCR products from EIEC strains.

To investigate the basis of this genetic variation, we cloned the 3.2 amplicons obtained from EIEC strain 1380, FG4 and the 3.9 kb amplicon

FG5 into the plasmid pGem-T. Sequence analysis of the resulting plasmids revealed that an IS*600* element (about 1.3 kb) is inserted into the *nadB* gene, after codon 52. At the insertion junction, we detected a duplication of a GTA sequence, present as direct repeats flanking IS*600* elements. The location and orientation of the IS*600* element were then confirmed by PCR analysis with a primer pair internal to the *nadB* gene and to the IS*600* sequence, respectively (Table 5, Figure 15). The sequence analysis of the *nadB* gene from *EIEC* FG5 reveals that the gene has been inactivated by the insertion of an IS*600* element, which carries, within its sequence, another *IS* element (IS*1*). A gc direct repeat is present at the IS*600*-IS*1* junction. In these EIEC strains IS disruptions of *nadB* is responsible for nicotinic acid auxotrophy by preventing the production of functional protein. An IS*600* in the same position has been observed also in other two EIEC strains present in the database: *E. coli* EDL1284 (accession number EF473674) and *E. coli* strain 1 (accession number EF473673).

In EIEC 4608, one of the four nicotinic auxotrofic strains, we identified not only the polymorphic change found in the other EIEC strains that are nicotinic acid independent but also three specific aminoacid change (Gly74Glu, Pro272Ser and Pro533Leu, 6). Since one of these change is probably responsible of the *nadB* inactivation, in order to solve this issue we performed in vivo complementation experiment using recombinant plasmid contained *E. coli nadB* gene from EIEC 4608 ($pNadB_{4608}$) in which the mutated aminoacid residue has been changed by direct site directed mutagenesis in wild type residue (Figure 12). The plasmid $pNadB_{4608}$ in which we replaced the GAG codon at position 74, encoding a glutamate residue, with a GGG codon encoding a glycine was used to complement *E. coli* JW2558 *nadB* mutant in minimal medium. Since this plasmid, named $pNadB_{4608gly74}$, was able to complement JW2558 (Table 2), we suppose that the glicine at position 74 is required for full activity of the L-aspartate oxidase. In addition we change the serine at position 272 (TCG codon) in wild type proline (CCG codon) in both $pNadB_{4608}$ and $pNadB_{4608gly74}$ to construct $pNadB_{4608pro272}$ and $pNadB_{4608gly74pro272}$ respectively. The results obtained using wild type residue *vs* mutated NadB protein of EIEC 4608 in complementation assay (Table 2) confirm that only the change of glycine at position 74 is responsible of NadB inactivation in this EIEC strains.

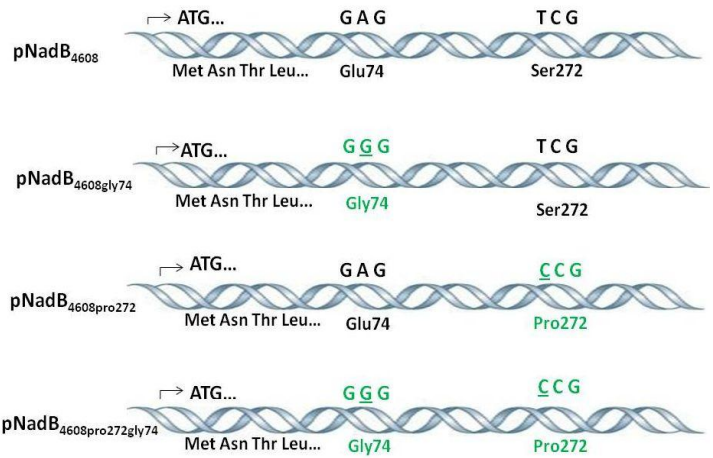

Figure 12: Site-specific mutagenesis of the codons in the ORF of *nadB* gene from EIEC 4608 cloned on a multycopy plasmid.

For that concern *nadA* genes of EIEC strains, we cloned and sequenced the PCR–generate *nadA* fragment from each of the them and perform DNA translation to protein and comparative alignment. The five EIEC strains serotype 0135 isolated in Somalia and 6.81 serotype O115 displayed no alterations of quinolate synthetase (NadA) compared to *E. coli* MG1655 (Table 5), the other strains showed one or more amino acids changes in NadA which are no effect on the functionality of the enzyme since EIEC FG1, FG2, FG6, FG7, 53638 are prototrophs and *nadA* gene of 4608, 1380, FG4 and FG5 is able to complement *E. coli nad*A mutant in minimal medium (Table 2).

Despite has been demonstrated that *Shigella* share nicotinic acid auxotrophy, our results demonstrated that nicotinic acid requirement is observed less frequently in enteroinvasive *E. coli* than in *Shigella.* Moreover all EIEC had no significant alteration of *nadA* gene and produced a functional quinolate synthetase, while the *nadB* gene is the preferential target for mutation resulting in nicotinic acid auxotrophy.

Table 2. Result of *nadB/nadA* complementation in EIEC

Strains	Grow in MM
MG1655	+
EIEC 4608	-
EIEC 4608 $pNadB_{mg}$	+
EIEC 1380	-
EIEC 1380 $pNadB_{mg}$	+
EIEC FG4	-
EIEC FG4 $pNadB_{mg}$	+
EIEC FG5	-
EIEC FG5 $pNadB_{mg}$	+
JW0733 *E. coli nadA* mutant	-
JW0733 $pNadA_{mg}$	+
JW0733 $pNadA_{4608}$	+
JW0733 $pNadA_{1380}$	+
JW0733 $pNadA_{FG4}$	+
JW0733 $pNadA_{FG5}$	+
JW2558 *E. coli nadB* mutant	-
JW2558 $pNadB_{mg}$	+
JW2558 $pNadB_{4608}$	-
JW2558 $pNadB_{4608gly74}$	+
JW2558 $pNadB_{4608pro\ 272}$	-
JW2558 $pNadB_{4608gly74\ pro\ 272}$	+

The subscript indicates the strain origin of *nadA* or *nadB* and the generated aminoacid change. $pNadA_{mg}$ and $pNadB_{mg}$ carry respectively the functional *nadA* and *nadB* gene from *E. coli* K-12 MG1655.

2.2 NICOTINIC ACID REQUIREMENT OF *SHIGELLA*

There are several well–known auxotrophic requirements of *Shigella* that are not found among most isolates of *E. coli*. Among these the nicotinic acid supplementation for growth on minimal medium due to mutation in one or both unlinked loci *nadA* and *nadB,* encoding the enzyme complex that converts L-aspartate to quinolate **(**Mantis 1996, Prunier *et al.,* 2007a). Recently as been reported that QUIN is a strong and specific inhibitor of several virulence phenotype of *Shigella* (Prunier *et al.,* 2007a). An in silico analysis performed on genome sequences currently available on public databases (http://www.mgc.ac.cn/ShiBASE/and http://www .ncbi.nlm .nih. gov/) highlights that *nadB* and/or *nadA* are always defective in *Shigella* and that its inactivation has been obtained by diverse strategies.

We analyse a large collection of *Shigella* strains isolated over several years in different geographic areas to respect the acid nicotinic requirement. All *Shigella* (Table 1) examined, required nicotinic acid or the intermediate quinolinic acid to grow suggesting a deficiency in *nadA* or *nad*B locus. Than to analyse the molecular changes responsible of *nadB* or *nadA* silencing we cloned and sequenced the *nadB* or *nadA* genes of 12 *S. flexneri* strains belong different serotype (serotype 1-6), two *S sonnei* and one *S dysenteriae* and tested their ability to restore the nicotinic acid prototrophy in *E. coli nadA* or *nadB* mutants, JW0733 and JW2558 respectively (Baba et al., 2006).

2.2.1 Molecular characterization of the quinolate synthetase in *Shigella*

In all the *S flexneri* amplification of the entire *nadA* gene, encoding for the quinolate synthetase, gave a product of the same size (1.4Kb) of the control strain MG1655, but only *nadA* from *S. flexneri*, serotype 3a (SFZM49, SFZM7), serotype 5 (M90T) and serotype 6 (SFZM2, SFZM9, SFZM43) was able to complement *E coli* JW0733 *nadA* mutant in minimal medium (Table 3). Than we sequenced *nadA* genes of these *S. flexneri* strains and compared each sequence to the translated *E. coli* K-12 sequence. This comparison (Table 4) shown that *nadA* gene from *S. flexneri* of these specific serotype contained point mutations that result in few or not amino-acid substitution. For that concern the NadA of the three *S. flexneri* strains serotype 6 (SFZM2, SFZM9 and SFZM43) we observed no changes compared to *E. coli* K-12 (Table 4), while *S. flexneri* serotype 3a (SFZM7and SFZM49) and serotype 5 (M90T) contained point mutations that result in two amino-acid substitution: ala111val and thr252ala (Table 4). These substitutions are not essential for the functionality of the quinolate synthetase since are able to complement JW0733 *nadA* mutant and are present also in EIEC strains FG1 that is prototroph (Table 4). These aminoacid changes are identical to those from the

published sequences for *S. flexneri* of the serotype 5a and 5b (Nie *et al.*, 2006; Onodera *et al.*, 2012) and *S. flexneri* serotype 3a available on public databases. To verify the functionality of quinolate synthase into these *Shigella* strains we introduced the pNadB_{mg} plasmid, containing the *nadB* gene from MG1655, into *S. flexneri* serotype 3a (SFZM7 and SFZM49), serotype 5 (M90T) and serotype 6 (SFZM2, SFZM9, SFZM43) to tested its ability to restore nicotinic acid independence. Since the introduction of a functional *nadB* restore the nicotinic acid independence only in *S. flexneri* M90T (Table 3) the most likely explanation for the absence of quinolate synthase activity in *S. flexneri* serotype 3a and serotype 6 includes the possibility of mutations in other regulators mapping outside the *nadA* gene. Alternatively, the absence of quinolate synthase activity may be due to mutations in other gene that impair the functionality of the quinolate synthase.

Recently has been demonstrated that the inactivation of the cysteine desulfurase IscS, involved in the synthesis and repair of the Fe–S cluster, produces severe growth defects such as nicotinic acid auxotrophy (Loiseau *et al.;* 2005). Since NadA contains a [4Fe-4S] which was shown to be essential for its activity, it is possible speculate that *S. flexneri* serotype 3a and serotype 6 are defective in post-translational mechanism involved in the synthesis of a functional [4Fe-4S] holoform of NadA. In *E. coli* the cysteine desulfurase IscS is encoded by the *iscRSUA* operon required for assembly of a variety of Fe–S enzymes In order to analyse if *iscS* defect are responsible of nicotinic acid auxotrophy we sequenced the entire *iscRSUA* operon of *S. flexneri* 3a for comparison with *E. coli* K-12 MG1655. Sequence analysis showed that there are no differences in the coding sequence, but highlighted the presence of a 183 base pairs deletion upstream the operon. Then to verify if the region deleted may be important for the expression of *iscRSUA* operon we constructed transcriptional fusions using *lacZ* as reporter gene. The results obtained by analyzing the expression of β-Galactosidase activity of plasmid PiscR_{MG}-lacZ (containing the region upstream *iscRSUA* operon of MG1655) and plasmid PiscR_{SFZM7}-lacZ (containing the region upstream *iscRSUA* operon, with the 187 bp deletion, of *S. flexneri* 3a strain SFZM7) show no difference in the β-Galactosidase activity. In addition the introduction of recombinant plasmid containing the entire *iscRSUA* operon of *E. coli* MG1655 (pIscRSUA NadB_{mg}) is unable to restore the nicotinic independence. These results indicated that *iscRSUA* system are not involved in the nicotinic acid requirement of *S. flexneri 3a*

Table 3. Complementation experiments of *nadB/nadA* from *Shigella*

Strains	Grow in MM
MG1655	+
JW0733 *E. coli nadA* mutant	-
JW0733 pNadA$_{mg}$	+
JW0733 pNadA$_{M90T}$	+
JW0733 pNadA $_{SFZM7/\ SFZM49}$	+
JW0733 pNadA$_{SFZM2/SFZM9/SFZM43}$	+
JW0733 pNadA $_{SFZM46\ YSH6000}$	-
JW0733 pNadA $_{SFZM8\ SFZM15\ SFZM53}$	-
JW0733 pNadA $_{SFZM50}$	-
JW0733pNadA $_{SFZM50\ gly304}$	-
JW0733 pNadA $_{SFZM50cys128}$	+
JW0733 pNadA $_{SDZM603}$	-
S.flexneri M90T pNadB$_{mg}$	+
S.flexneri SFZM7, SFZM49 pNadB$_{mg}$	-
S.flexneri SFZM2, SFZM9, SFZM43 pNadB$_{mg}$	-
S. dysenteriae SDZM603 pNadB$_{mg}$	-
JW2558 *E. coli nadB* mutant	-

The subscript indicates the strain origin of *nadA* or *nadB* and the generated aminoacid change. pNadA$_{mg}$ and pNadB$_{mg}$ carry respectively the functional *nadA* and *nadB* genefrom *E. coli* K-12 MG1655.

The *nadA* gene of *S. flexneri* serotype 1b (SFZM50), serotype 2a (SFZM46 and YSH6000), serotype 4 (SFZM8, SFZM15 and SFZM53) which was unable to complement *E.coli* JW0733 (Table3) in addition to change observed in *S. flexneri* containing a functional *nadA* gene showed two aminoacid change: cysteine at position 128 in tirosine (C128Y) and glycine at position 304 in aspartic acid (G304D) (Table 4) that may be responsible for the loss of functionality of the quinolate synthetase. To investigate if one or both changes are responsible of NadA inactivation we performed in vivo complementation experiment using recombinant plasmid contained *nadA* gene from SFZM50 (pNadA$_{SFZM50}$) in which the single aminoacid residue has been changed by direct site directed mutagenesis (Figure 13) in wild type residue. Only the plasmid pNadA$_{SFZM50cys128}$ in which we replaced the TAC codon at position128, encoding a tyrosine residue, with a TGC codon encoding a cisteine was able to complement *E. coli* JW0733 *nadB* mutant in minimal medium (Table3). In addition we change the 304 aspartic acid residue in wild type glycine to construct the plasmid pNadA$_{SFZM50gly304}$ enable to restore nicotinic independence in *E. coli* JW0733. This result confirm that

only the change C128Y is responsible of quinolate synthase inactivation (NadA) in *S flexneri* SFZM50 and in other *S flexneri* (SFZM46 YSH6000, SFZM8, SFZM15 SFZM53) that show the same aminoacid change (serotype 4). Since in *E. coli*, the prototype for the study of quinolate synthase, as been demonstrated that the cysteine at position 113, 200 291 294 297 are important for the activity of the enzyme and that serine substitution at cysteine 128 had no effect on NadA functionality (Saunders *et al.*, 2008) **t**he result that we obtained are consisted with the effect of a severe change of cisteine in tyrosine.

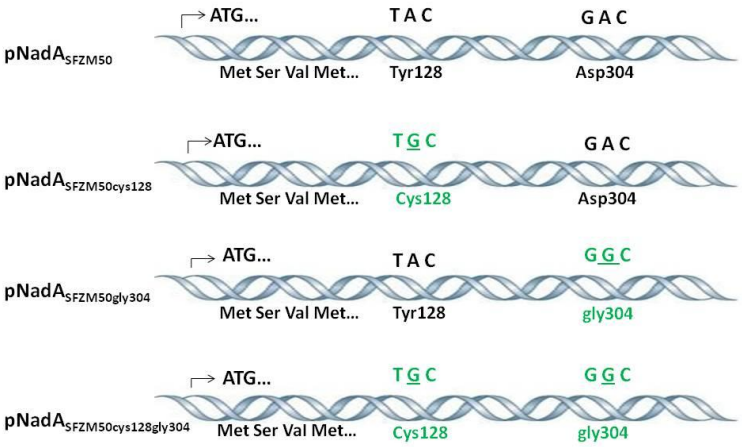

Figure 13. Site-specific mutagenesis of the codons in the ORF of *nadA* gene from SFZM50 cloned on a multycopy plasmid.

Amplification of *nadA* gene from the two *S. sonnei* strains resulted in a product larger about 2.1Kb than the control *E. coli* K-12 MG1655. Sequences analysis confirmed that in *S. sonnei nadA* gene showed disruption by an IS*21 between* codon 292 and 293.

The *nadA* gene of *S. dysenteriae1* (SDZM603) was unable to complement *E. coli* JW0733 (Table 3). The sequence analysis and comparison with *S. dysenteriae*197(accession number NC00760) reveals that both strains displayed the same amino acid changes in NadA.

Taken together these results suggest that in *Shigella* inactivation of quinolate-synthetase is an event that occurs through point mutations and IS

interruption of the *nadA* gene or defect in the machinery involved in the synthesis of the correct holoform of the protein.

2.2.2 Molecular characterization of the L-aspartate oxidase in *Shigella*

Although all *Shigella flexneri* studied, except *S flexneri* of serotype 6, gave a *nadB* amplicon of the same size (1.9Kb) of the control strain MG1655 (Figure 14), only *S flexneri* SFZM7 and SFZM49 (serotype 3a) contained a functional *nadB*, able to complement *E.coli* JW2558 *nadB* mutant (Table 3). Moreover the aminoacid changes found in NadB sequence of *S. flexneri* serotype 3a were present also in other strains, as well as EIEC strains, with functional NadB protein (Table 5).

S flexneri serotype 6 (SFZM2, SFZM9, SFZM46) gave a product 1.3 kb larger of the control strain, cloning and sequence analysis of the resulting plasmids ($pNadB_{SFZM2}$, $pNadB_{SFZM9}$, $pNadB_{SFZM46}$) revealed that the *nadB* gene in these strains is disrupted by insertion of an IS*600* element (about 1.3 kb) between codon 178 and 179. At the insertion junction, we detected a duplication of a GGG sequence, present as direct repeats flanking IS*600* elements. Interestingly an IS*600* in the same position has been observed in *S .boydii* 227 (accession number NC007613).

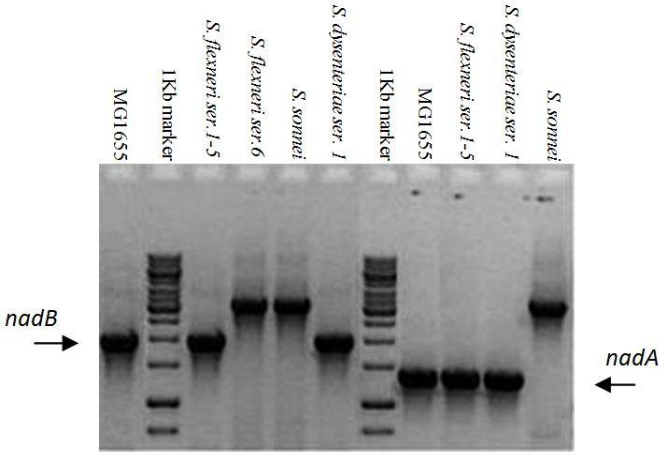

Figure 14. Agarose-gel electrophoresis of the *nadA* and *nadB* PCR products obtained from *E. coli* K12 MG1655 and *Shigella*.

Moreover *S. flexneri* serotype 1b (SFZM50), 2a (SFZM46, YSH6000), 4a (SFZM8, SFZM15, SFZM53) and 5a (M90T) failed to grown on minimal medium when transformed whit *nadA* functional gene indicating that NadB protein was unfunctional (Table 3). Sequence analysis revealed that in these strains *nadB* inactivation is due to point mutation that created a TGA stop codon at amino acid 354 (Table 5), resulting in truncation of the C-terminal third of the *nadB* product. (Figure 16). The presence of the same mutation at position 354 of the peptide was confirmed in all *S. flexneri* serotype 1, 2, 4 and 5 used in this study such as in other *S. flexneri* of these specific serotype.
In *S. sonnei* SSZM34 and SSZM45 the *nadB* gene is disrupted by insertion of an IS*600* element in codon 233. Their sequences are identical to the *nadA* and *nadB* loci of *S. sonnei* 46, the sequenced strain from the databases. These results are consistent with the well known clonal nature of *S. sonnei* strains. In all of these strains, it is very likely that the IS-related disruptions of *nadA* and/or *nadB* loci are responsible for the nicotinic acid auxotrophy by preventing the production of functional proteins.

Inactivation of L-aspartate oxidase, the first enzyme steps in the pathway for the *de novo* synthesis of NAD, is an event that occurs frequently in *Shigella* through several strategy which may include point mutations or insertion of IS element. In this context an exception is represented by *S. flexneri* serotype 3a containing a functional NadB.

Few aminoacid changes are found in SFZM7 and SFZM49 (both serotype 3a) compared to *E. coli*-K12 that are not responsible for the nicotinic acid requirement, since *nadB* from SFZM7 and SFZM49 is able to complement *nadB* mutant of *E. coli* K-12. Moreover the aminoacid changes found in NadB sequence of *S. flexneri* serotype 3a were present also in other strains of *S. flexneri* serotype 3a, as well as EIEC strains, with functional NadB protein Despite their inability to synthesize nicotinic acid the *S flexneri* serotype 3a encode both NadA and NadB able to complement *E. coli* K-12 *nadA* and *nadB* mutants respectively. Taken together, these results suggest that other genes are inactive in these specific serotype.

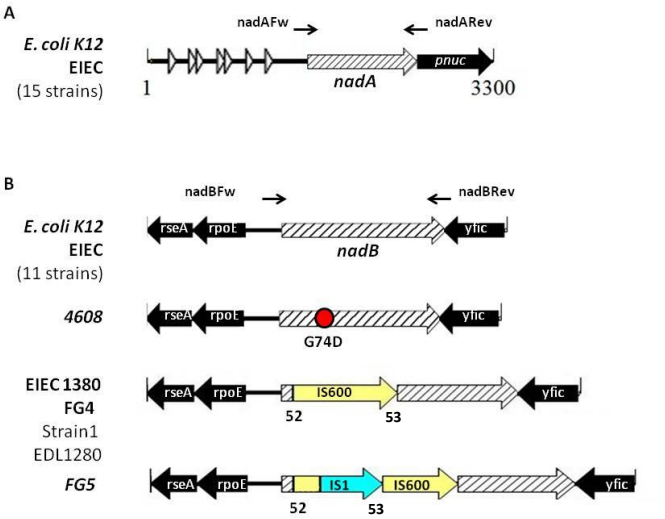

Figure 15. Comparative analysis of the *nadA* (A) and *nadB*(B) loci of EIEC and *E. coli* K-12. Arrows indicate the orientation of the IS sequences and of the *nadB* open reading frames. The locations of the primers used are shown below the appropriate region. The positions of IS and codon change were confirmed by sequence analysis.

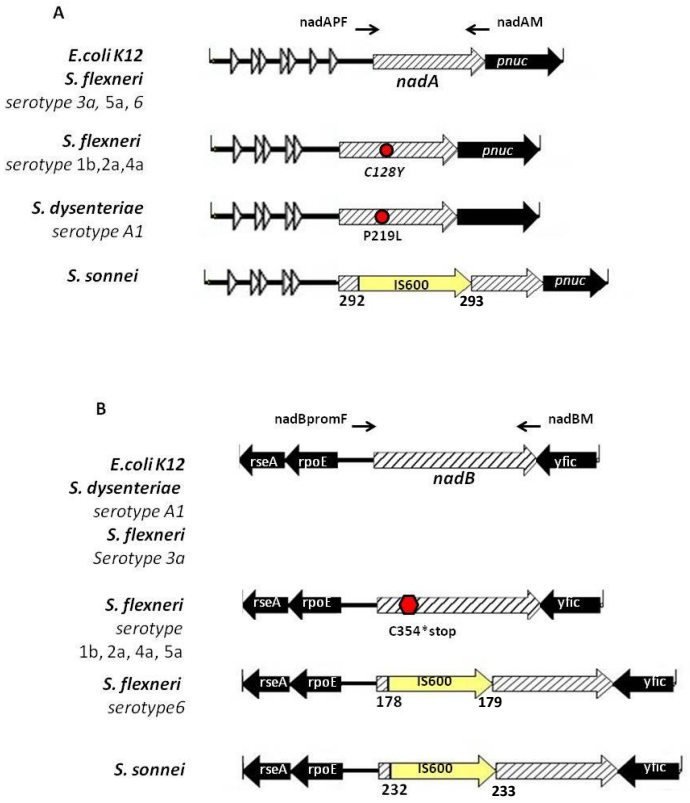

Figure 16. Comparative analysis of the *nadA* (A) and *nadB* (B) loci of *Shigella* and *E. coli* K-12. Arrows indicate the orientation of the IS sequences and of the *nadB* open reading frames. The locations of the primers used are shown below the appropriate region. The positions of IS and codon change were confirmed by sequence analysis.

Table 4: NadA aminoacid changes in EIEC and *S.flexneri* strains compared to *E. coli* K-12 MG1655

Strains	**Aminoacid changes**																	
MG1655	**102 K**	**111 A**	**128 C**	**130 A**	**134 R**	**145 A**	**148 A**	**160 V**	**191 G**	**198 G**	**219 P**	**252 T**	**253 L**	**257 R**	**292 R**	**304 G**	**324 D**	**343 A**
HN11,HN13, HN19, HN280, HN300																		
FG1	R	V		T								A						
FG2, FG6, 53638																	E	
FG7									A				S	W				
FG4, FG5, 1380										S								
4608					H			A	A									
SFZM2, SFZM9 SFZM43																		
M90T		V										A						
SFZM7 SFZM49		V					T					A						V
SFZM50 SFZM46,YSH6000		V	Y									A				D		
SFZM8,SFZM15;SFZM53		V	Y			D						A				D		
SSZM34, SSZM45															IS			
SDZM608					H				A		L			W				

Table 5: NadB aminoacid changes in EIEC and *S.flexneri* strains compared to *E. coli* K-12 MG1655

Strains	Aminoacid changes																				
MG1655	52 I	73 A	74 G	80 R	95 Q	102 V	108 I	112 G	141 142 149 E T L	167 V	178 V	191 T	219 P	232 G	272 P	283 E	354 C	412 415 416 Y D I	468 T
HN11 HN13 HN19 HN280 HN300										I I I I I		K K K K K				V			
6.81								S		I						V		G V	N
FG1							V		Q S Q									G V	
FG2										I						V			
FG6										I		K				V			
53638						G				I		K				V			
FG7							V		Q S Q									D G V	
FG4 FG5 1380	*IS*						V		Q S Q									D G V	
4608			E							I					S	V		G V	N
SFZM7 SFZM49		S					V		Q S Q									G V	
SFZM50 SFZM46 YSH6000 M90T				H	P		V		Q S Q								stop	G V	
SFZM8 SFZM15 SFZM53				H	P		V		Q S Q								stop	G V	
SFZM2 SFZM9 SFZM43							V		Q S Q		*IS*							D G V	N
SSZM34														*IS*					
SDZM608										I			L					V	

2.3 ANALYSIS OF QUIN INFLUENCE ON VIRULENCE GENES EXPRESSION IN *SHIGELLA*/EIEC STRAINS.

Recently as been reported that QUIN inhibits several virulence phenotypes of *Shigella* (Prunier *et al.*, 2007a), in particular bacterial invasion, cell-to-cell spread and induction of PMN transepithelial migration by interfering with Type Three Secretion System (T3SS) which are the key to the entry of the bacteria into eukaryotic cells providing a pore for the delivery of many effectors. Has been demonstrated that QUIN shows a strong inhibitory effect on the induction of TTSS, as shown by reduced levels of secreted Ipa proteins when the compound is added during the secretion assay (Prunier *et al*., 2007a).

To evaluate if the target of QUIN could be the expression of the virulence genes, we used gene fusions in which the gene for β-galactosidase was placed under the control of the promoters of virulence genes *mixi*, *mixC* and *ipaB*. These fusions were transferred by P1 transduction from *Shigella flexneri* 2a strains BS184, BS228 and BS230 in EIEC HN570 (Δlac derivative of EIEC HN280) to construct HN184, HN228 and HN230 respectively. The activity of the three different promoters were measured in presence or absence of 1mM QUIN.

However, in *Shigella* and EIEC strains, QUIN reduces the activity of the promoters tested (Table 6).

Table 6. Expression of β-galactosidase

Strain	Relevant Genotype	β–Galactosidase activity	
		LB	LB+Quin
BS184	*S. flexneri mxiC*::*lacZ*	245	153
BS228	*S. flexneri ipaB*::*lacZ*	428	158
BS230	*S. flexneri mxi*::*lacZ*	663	394
HN184	HN570 *mxiC*::*lacZ*	156	81
HN228	HN570 *ipaB*::*lacZ*	209	102
HN230	HN570 *mxi*::*lacZ*	274	135

[a]Units of β-galactosidase are calculated according to Miller (1992).
The results represent the average of at least four independent experiments.

These results confirm that a bacterial target for the action of QUIN is probably the transcription of the gene encodes the components needed for the assembly and function of the T3SS, and in particular of the effectors. This assay also showed that the activity of the promoters analyzed was lower in EIEC than in *Shigella*, suggesting a difference in the expression of the virulence phenotype between these two bacteria.

DISCUSSION

In bacteria the evolution toward pathogenic phenotypes revolves around two distinct mechanisms, the acquisition of additional genes encoding virulence determinants and the loss or modification of preexisting genetic material. The acquisition of multiple linked virulence traits by horizontal gene transfer is very important in triggering a virulence phenotype in a commensal organism,which can thus gain access to new host environments. The new pathogen then reaches optimal fitness within the novel environment by adaptation of its genome through mutations in preexisting genes. These so-called pathoadaptive mutations, in gene defined antivirulence loci(AVL), that improve survival within host tissues, increase the pathogenic potential of the bacteria, and drive the evolution of a microrganism toward a more pathogenic phenotype (Casalino *et al.,* 2003)

In this context, evolution of *Escherichia coli* toward pathogenicity represents an interesting model for the enormous versatility of this microrganism, causing an impressive variety of different diseases. Acquisition of virulence genes by horizontal transfer has played an important role in the evolution of pathogenic *E. coli* strains. Virulence determinants have been acquired by *E. coli* cells as parts of plasmids, bacteriophages, transposons, or pathogenicity islands

Enteroinvasive *E. coli* (EIEC) strains are facultative intracellular pathogens able to enter epithelial cells of the colon, replicate within them, and move between adjacent cells with a mechanism similar to that of *Shigella*, the causative agent of bacillary dysentery. In EIEC and *Shigella*, the critical event in the transition toward a pathogenic life-style probably relies on the acquisition of a large (220- to 260-kb) F-type plasmid (pINV) This plasmid contains all the genes required for invasion and for intra- and intercellular spread, including their positive activators. Plasmid virulence genes are silenced outside the host by one of the major histone-like proteins, H-NS, which acts as a repressor on two positive activators. pINVs isolated from EIEC and *Shigella* strains share wide regions of high structural and functional homology.

Although EIEC strains may be developing the full *Shigella* phenotype, they do not have the full set of characters that define *Shigella* strains and are not included in any of the three *Shigella* clusters. EIEC strains generally correspond to bioserotypes found in a dozen *E. coli* serogroups, and recent molecular analyses confirm that they are widely distributed among *E. coli* phylogenetic groups (Pupo *et al*., 2000).

Many EIEC strains have *Shigella*-like features: they can be Lac^{-}; nonmotile, low-level indole producing, and/or unable to produce gas during fermentation. Some have the same O antigen that is present in *Shigella*. In contrast to *Shigella*, they have a high metabolic activity since they still retain the ability to catabolize substrates widely utilized by *E. coli*.

A characteristic of *Shigella*, is their strict requirement for an exogenous source of nicotinic acid in that the bacterium is unable to synthesize NAD via the *de novo* pathway but can still synthesize NAD from exogenous nicotinic acid. Mutations in *nadA* and *nadB* are responsible for this auxotrophy in *S. flexneri 2a* (Gemski *et al*., 1971). Recently has been proposed that silencing of the *nadA or nadB loci* in *Shigella* represents an important pathoadaptive mutation necessary for increasing the pathogenic potential of bacteria in host tissues since the quinolinate, an intermediate in the *de novo* pathway of NAD cause a strong attenuation of virulence of *S. flexneri* 5a (Prunier *et al.*, 2007a).

The difference in the ability to synthesize NAD via the *de novo* pathway between *Shigella* and EIEC O135 isolated in Somalia was our *rationale* for a comparison of *nadA* and *nadB* genes between *Shigella* and EIEC. In order to evaluate if the auxotrophic requirement for nicotinic acid may be selected also in EIEC strains, which are as regarded as evolutionary intermediates from *E .coli* to *Shigella* we analyse several EIEC and *Shigella* strains, isolated in different geographic areas, belong to different serotypes.

Our observation in 15 EIEC strains reveal that nicotinic acid requirement is observed less frequently in enteroinvasive *E. coli* than in *Shigella.* Moreover all EIEC had no significant alteration of *nadA* gene and produced a functional quinolate synthetase (NadA), since the introduction of a functional *nadB* gene restores wild-type activity. In few EIEC strains *nadB* gene is the target for mutations resulting in nicotinic acid auxotrophy and the main strategy adopted for silencing the *nad*B gene is disruption by IS element. Interestingly, in the *nadB* gene of two EIEC strains, present in databases as been observed the same IS element in the same position. Since the majority of EIEC strains examined showed a NAD biosynthetic pathway and seems to be more similar to *E. coli* than *Shigella*, EIEC could be considered a midpoint in the evolution between *E. coli* and *Shigella*

All Shigella strains examined (12 *S flexneri* serotype 1-6, one *S dysenteriae 1* and two *S. sonnei*) required nicotinic acid or the intermediate quinolinic acid suggesting a deficiency in *nadA* or *nad*B locus. In order to determine if some of these changes were due to polymorphisms of the *nadA* and *nadB* loci, we cloned and sequenced

the *nadB* or *nadA* genes of all *Shigella* and perform DNA translation to protein for comparative alignment with sequences of *E. coli* and *Shigella* available in the databases. Combining result from genetic approach and comparative genomic analyses we observe that in *S. flexneri* despite the difference in geographic origin and antibiotic resistance patterns, the strains of the same serotype have undergone the same modifications in both the nadA and nadB genes.

All *Shigella flexneri*, except *S. flexneri* serotype 3a contained alterations in the *nadB* gene encoding for L-aspartate oxidase, the first enzyme steps in the pathway for *de novo* synthesis of NAD. This is an event that occurs through two strategies: point mutations leading to the replacement of an aminoacid by a stop codon, (*S. flexneri* serotype 1b, 2a, 4a and 5a) or IS-related disruption responsible for the nicotinic acid auxotrophy by preventing the production of functional proteins (*S. flexneri* serotype 6) Interestingly an *IS* element in the same position has been observed in *S. boydii* and on genome sequences currently available on public databases of *S. flexneri* serotype 6 and *S. sonnei.*

Inactivation of quinolate synthase (NadA) occurs through a change of cysteine 128 in tyrosine (*S. flexneri* serotype 1b, 2a and 4a). *S. flexneri* serotype 3a, 5 and 6 contains a functional *nadA* gene able to complement *E. coli nadA* mutant. Since the introduction of a functional *nadB* restore the nicotinic acid independence in M90T (*S. flexneri* serotype 5) this result is in accord with Mantis *et al.* that in 1996 reported defects limited to *nadB* in M90T. On the contrary the introduction of a functional *nadB* is unable to restore the nicotinic acid independence in *S. flexneri* serotype 3a and 6; the most likely explanation for the absence of quinolate synthase activity in these *S. flexneri* includes the possibility of mutations mapping outside the *nadA* gene that impair the functionality of quinolate synthase in these specific serotype.

Sequences analysis reveals that in the two *S. sonnei examined* both the *nadA* and *nadB* loci are disrupted by IS element, inserted in the same position of other *S sonnei* present in data base respectively. These results are consistent with the well known clonal nature of *S. sonnei* strains.

Several of our observations suggest that mutation of NadB was the first event in the evolution of *Shigella* species toward nicotinic acid auxotrophy. This phenomenon could be explained by the larger size of the *nadB* locus relative to *nadA*, which makes it a more frequent target for random mutation during the selection process for nicotinic acid auxotrophy. The loss of *nadA* functions could be the result of genetic drift of *nadA* after inactivation of *nadB* as a result of the loss of the

selective pressure to maintain a functional *nadA* gene In fact EIEC and the majority of *Shigella* strains studied show no modifications in NadA. Interestingly *S.flexneri* serotype 3a, represent an exception since is not able to grow in minimal while presenting both *nadB* and *nadA.* functional genes.

Our observations are consistent with previous report that the *nadA* and *nadB* genes are AVL in *Shigella* (Prunier *et al.*, 2007a). Additionally, the distinct nature of mutations defining the *nadA* and *nadB* loci of the four different species of *Shigella* suggests their acquisition and accumulation after *Shigella* speciation. This conclusion is in accord with the model that clones of *Shigella* evolved multiple times from different lines of *E. coli* ancestor strains. Our findings confirm and extend the concept that *Shigella* evolution proceeds through convergent evolution toward removal or inactivation of AVL by whatever alteration (IS, deletion, or point mutation) leads to a more virulent phenotype or to improved fitness for survival in the host.

Although EIEC strains may be developing the full *Shigella* phenotype, they do not have the full set of characters that define Shigella strains and are not included in any of the three *Shigella* clusters. The majority of the EIEC analyzed in this study are nicotinic acid independent and in the auxotrophs strains the defect is limited to *nadB*. Since EIEC retained more characteristics of commensal *E. coli* than *Shigella* spp. such as nicotinic acid indipendence, these strains might reflect an earlier stage of the evolutionary process undergone by *Shigella*

REFERENCES

Barnhart, M.M., Chapman, M.R. 2006. Curli biogenesis and function. Annu. Rev. Microbiol. **60**, 131e147.

Barbagallo, M., Di Martino, M.L., Marcocci, L., Pietrangeli, P., De Carolis, E., Casalino, M., Colonna, B., Prosseda, G. 2011. A new piece of the Shigella pathogenicity puzzle: spermidine accumulation by silencing of the speG gene. PLoS One **6**, e27226.

Bernardini, M. L., J. Mounier, H. d'Hauteville, M. Coquis-Rondon, and P. J. Sansonetti. 1989. Identification of *ics*A, a plasmid locus of *Shigella flexneri* that governs bacterial intra- and intercellular spread through interaction with F-actin. Proc. Natl. Acad. Sci. USA **86:**3867–3871.

Blocker, A., P. Gounon, E. Larquet, K. Niebuhr, V. Cabiaux, C. Parsot, andP. Sansonetti. 1999. The tripartite type III secreton of *Shigella flexneri*inserts IpaB and IpaC into host membranes. J. Cell Biol. **147:**683–693.

Brussow H, Canchaya C & Hardt WD. 2004. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. Microbiol Mol Biol Rev **68**: 560–602

Buchrieser, C., P. Glaser, C. Rusniok, H. Nedjari, H. D'Hauteville, F.Kunst, P. Sansonetti, and C. Parsot. 2000. The virulence plasmid pWR100 and the repertoire of proteins secreted by the type III secretion apparatus of *Shigella flexneri*. Mol. Microbiol. **38:**760–771.

Buysse, J. M., C. K. Stover, E. V. Oaks, M. Venkatesan, and D. J. Kopecko.1987.Molecular cloning of *invasion plasmid antigen* (*ipa*) genes from *Shigellaflexneri*: analysis of *ipa* gene products and genetic mapping. J. Bacteriol. **169:**2561–2569.

Casalino M., Yusuf M. W., Nicoletti M., Bazzigalupo P., Coppo A., Colonna B., Cappelli C., Bianchini C., Falbo V., Ahmed H. J., Omar K. H., Maxamuud K. B., Maimone F. 1988. A two-year study of enteric infections associated with diarrhoeal diseases in children in urban Somalia. Trans. R. Soc. Trop. Med. Hyg. 82(4):637-641.

Casalino,M., Latella,M.C., Prosseda,G., and Colonna,B. 2003 CadC is the preferential target of a convergent evolution driving enteroinvasive *Escherichia coli* toward a lysine decarboxylase-defective phenotype. Infect. Immun. 71, 5472-5479.

Casalino,M., Latella,M.C., Prosseda,G., Ceccarini,P. Grimont,F. and Colonna,B. 2005. Molecular evolution of the lysine decarboxylase-defective phenotype in *Shigella sonnei.* Int J Med Microbiol. 2005 Mar;294(8):503-12.

Casalino, M., Prosseda, G., Barbagallo, M., Iacobino, A., Ceccarini, P., Latella, M.C., Nicoletti, M., Colonna, B. 2010. Interference of the CadC regulator in the arginine-dependent acid resistance system of Shigella and enteroinvasive E. coli. Int. J. Med. Microbiol. **300**, 289-295.

Casalino, M., M. W. Yusuf, M. Nicoletti, P. Bazzicalupo, A. Coppo, B. Colonna, C. Cappelli, C. Bianchini, V. Falbo, H. J. Ahmed, K. H. Omar, K. B. Maxamuud, and F. Maimone. (1988). A two-year study of enteric infections associated with diarrhoeal diseases in children in urban Somalia. Trans. R. Soc. Trop. Med. Hyg. 82:637-641.

Day,W.A., Jr., Fernandez,R.E., and Maurelli,A.T. 2001. Pathoadaptive mutations that enhance virulence: genetic organization of the cadA regions of *Shigella* spp. Infect. Immun. 69, 7471-7480.

Dobrindt, U., B. Hochhut, U. Hentschel, and J. Hacker. 2004. Genomic islands in pathogenic and environmental microorganisms. Nat. Rev. Microbiol.**:**414–424.

Durand, J.M., Bjork, G.R. 2009. Metabolic control through ornithine and uracil of epithelial cell invasion by Shigella flexneri. Microbiology **155**, 2498-2508

Henderson, I. R., J. Czeczulin, C. Eslava, F. Noriega, and J. P. Nataro. 1999. Characterization of Pic, a secreted protease of *Shigella flexneri* and enteroaggregative *Escherichia coli*. Infect. Immun. **67:**5587–5596.

Escobar-Paramo P, Giudicelli C, Parsot C, Denamur E. 2003. The evolutionary history of *Shigella* and enteroinvasive *Escherichia coli* revised. J Mol Evol. ;57(2):140-8.

Fasano, A., F. R. Noriega, F. M. Liao, W. Wang, and M. M. Levine. 1997.Effect of *Shigella* enterotoxin 1 (ShET1) on rabbit intestine in vitro and in vivo. Gut **40:**505–511.

Fernandez,I.M., Silva,M., Schuch,R., Walker,W.A., Siber,A.M., Maurelli,A.T., and McCormick,B.A. 2001. Cadaverine prevents the escape of *Shigella flexneri* from the phagolysosome: a connection between bacterial dissemination and neutrophil transepithelial signaling. J. Infect. Dis. 184, 743-753.

Sokurenko, E. V., Hasty, D. L. & Dykhuizen, D. E. (1999). Pathoadaptive mutations: gene loss and variation in bacterial pathogens. Trends Microbiol 7, 191–195.

Flachmann, R., Kunz, N., Seifert, J., Gutlich, M.,Wientjes, F. J., Laufer, A. & Gassen, H. G. 1988. Molecular biology of pyridine nucleotide biosynthesis in *Escherichia coli*. Cloning and characterization of quinolinate synthesis genes *nadA* and *nadB*. Eur J Biochem **175**, 221–228

Formal, S.B., Gemski Jr., P., Baron, L.S., LaBrec, E.H. 1971. A chromosomal locus which controls the ability of Shigella flexneri to evoke keratoconjunctivitis. Infect. Immun. **3**, 73-79

Foster, J.W. 2004. Escherichia coli acid resistance: tales of an amateur acidophile. Nat. Rev. Microbiol. **2**, 898-907.

Foster, J. W., and A. G. Moat. 1980. Nicotinamide adenine dinucleotide biosynthesis and pyridine nucleotide cycle metabolism in microbial systems. Microbiol. Rev. **44**:83–105.

Fukushima, M., K. Kakinuma, and R. Kawaguchi. 2002. Phylogenetic analysis of *Salmonella*, *Shigella*, and *Escherichia coli* strains on the basis of the *gyrB* gene sequence. J. Clin. Microbiol. **40**:2779–2785.

Gemski, P., Formal, S. B. & Baron, L. S. 1971. Identification of two widely separated loci conferring nicotinic acid dependence on wildtype
Shigella flexneri 2a. Infect Immun 3, 500–503.

Griffith G.R. , Jerry LR Chandler, And Robert K. 1975. Gholson. Studies on the *de novo* Biosynthesis of NAD in *Escherichia coli.* The Separation of the *nadB* Gene Product from the *nadA* Gene Product and Its Purification. Eur. J. Biochem. **54**,239-245.

Groisman EA & Ochman H. 1996. Pathogenicity islands: bacterial evolution in quantum leaps. Cell **87**: 791–794.

Hacker J & Kaper JB. 2000. Pathogenicity islands and the evolution of microbes. Annu Rev Microbiol **54**: 641–679.

Hromockyj, A. E., and A. T. Maurelli. 1989. Identification of *Shigella* invasion genes by isolation of temperature-regulated *inv*::*lacZ* operon fusions. Infect. Immun. **57:**2963–2970.

Ingersoll, M., E. A. Groisman, and A. Zychlinsky. 2002. Pathogenicity islands of *Shigella*. Curr. Top. Microbiol. Immunol. **264:**49–65.

Ingersoll, M. A., J. E. Moss, Y. Weinrauch, P. E. Fisher, E. A. Groisman, and A. Zychlinsky. 2003. The ShiA protein encoded by the *Shigella flexneri* SHI-2 pathogenicity island attenuates inflammation. Cell. Microbiol. **5:**797– 807.

Ingersoll, M. A., and A. Zychlinsky. 2006. ShiA abrogates the innate T-cell response to *Shigella flexneri* infection. Infect. Immun. **74:**2317–2327.

Jiang, Y., F. Yang, X. Zhang, J. Yang, L. Chen, Y. Yan, H. Nie, Z. Xiong, J. Wang, J. Dong, Y. Xue, X. Xu, Y. Zhu, S. Chen, and Q. Jin. 2005. The complete sequence and analysis of the large virulence plasmid pSS of *Shigella sonnei*. Plasmid **54:**149–159.

Jin, Q., Z. Yuan, J. Xu, Y. Wang, Y. Shen, W. Lu, J. Wang, H. Liu, J. Yang,F. Yang, X. Zhang, J. Zhang, G. Yang, H. Wu, D. Qu, J. Dong, L. Sun, Y.Xue, A. Zhao, Y. Gao, J. Zhu, B. Kan, K. Ding, S. Chen, H. Cheng, Z. Yao,B. He, R. Chen, D. Ma, B. Qiang, Y. Wen, Y. Hou, and J. Yu. 2002. Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity through comparison with genomes of *Escherichia coli* K12 and O157. Nucleic Acids Res.**30:**4432–4441.

Kane, C. D., R. Schuch, W. A. Day, Jr., and A. T. Maurelli. 2002. MxiE regulates intracellular expression of factors secreted by the *Shigella flexneri* 2a type III secretion system. J. Bacteriol. **184:**4409–4419.

Kaper,J.B., Nataro,J.P., Moley,H.L. 2004. Pathogens *Escherichia coli.* Nat. Rev. Microbiol. Feb; (2)2: 123-140

Kotloff KL, Winickoff JP, Ivanoff B, Clemens JD, Swerdlow DL, Sansonetti PJ, Adak GK, Levine MM (1999) Global burden of *Shigella* infections: implications for vaccine development and implementation of control strategies. Bull WHO 77:651–666.

Lan R., Lumb B., Ryan D., Reeves P. R., 2001. Molecular evolution of large virulence plasmid in *Shigella* clones and enteroinvasive *Escherichia coli*. Infect. Immun. Oct;69(10):6303-9.

Lan, R., and P. R. Reeves. 2002. *Escherichia coli* in disguise: molecular origins of *Shigella*. Microbes Infect. **4:**1125–1132.

Lan R, Alles MC, Donohoe K, Martinez MB, Reeves PR. 2004. Molecular evolutionary relationships of enteroinvasive *Escherichia coli* and *Shigella* spp. Infect Immun **72**:5080–5088

Lindberg, A. A., A. Karnell, and A. Weintraub. 1991. The lipopolysaccharide of Shigella bacteria as a virulence factor. Rev. Infect. Dis. 13(Suppl. 4):S279–S284.

Lindsey,D.F., Mullin,D.A., and Walker,J.R. 1989. Characterization of the cryptic lambdoid prophage DLP12 of *Escherichia coli* and overlap of the DLP12 integrase gene with the tRNA gene *argU*. J. Bacteriol. 171, 6197-6205.

Mantis, N. J., and P. J. Sansonetti. 1996. The *nadB* gene of *Salmonella typhimurium* complements the nicotinic acid auxotrophy of *Shigella flexneri*. Mol. Gen. Genet. **252:**626–629.

Maurelli,A.T., Fernandez,R.E., Bloch,C.A., Rode,C.K., and Fasano,A. 1998. "Black holes" and bacterial pathogenicity: a large genomic deletion that enhances the virulence of *Shigella* spp. and enteroinvasive *Escherichia coli*. Proc. Natl. Acad. Sci. U. S. A 95, 3943-3948.

Maurelli AT. 2007. Black holes, antivirulence genes, and gene inactivation in the evolution of bacterial pathogens. FEMS Microbiol Lett.;267(1):1-8.

McCormick, B. A., A. M. Siber, and A. T. Maurelli. 1998. Requirement of the *Shigella flexneri* virulence plasmid in the ability to induce trafficking of neutrophils across polarized monolayers of the intestinal epithelium. Infect. Immun. **66:**4237–4243.

McCormick, B.A., Fernandez, M.I., Siber, A.M., Maurelli, A.T. 1999. Inhibition of Shigella flexneri-induced transepithelial migration of polymorphonuclear leucocytes by cadaverine. Cell. Microbiol. **1,** 143-155.

Mènard, R., P. J. Sansonetti, and C. Parsot. 1993. Nonpolar mutagenesis of the *ipa* genes defines IpaB, IpaC, and IpaD as effectors of *Shigella flexneri* entry into epithelial cells. J. Bacteriol. **175:**5899–5906.

Miller, J. H. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Moran, N. A. 2002. Microbial minimalism: genome reduction in bacterial pathogens. Cell **108:**583–586.

Miller, J. H. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Nakata, N., Tobe,T., Fukuda, I., Suzuki, T., Komatsu,K., Yoshikawa, M., and Sasakawa, C. 1993. The absence of a surface protease, OmpT, determines the intercellular spreading ability of *Shigella*: the relationship between the *ompT* and *kcpA* loci. Mol. Microbiol. 9, 459-468.

Nie, H., F. Yang, X. Zhang, J. Yang, L. Chen, J. Wang, Z. Xiong, J. Peng, L. Sun, J. Dong, Y. Xue, X. Xu, S. Chen, Z. Yao, Y. Shen, and Q. Jin. 2006. Complete genome sequence of *Shigella flexneri* 5b and comparison with *Shigella flexneri* 2a. BMC Genomics **7:**173.

Niyogi, S. K. 2005. Shigellosis. J. Microbiol. **43:**133–143.

Ochman, H., Moran, N.A., 2001. Genes lost and genes found: evolution of bacterial pathogenesis and symbiosis. Science 292, 1096e1098.

Page, A. L., H. Ohayon, P. J. Sansonetti, and C. Parsot. 1999. The secreted IpaB and IpaC invasins and their cytoplasmic chaperone IpgC are required for intercellular dissemination of *Shigella flexneri*. Cell. Microbiol. **1:**183–193.

Peng, J., X. Zhang, J. Yang, J. Wang, E. Yang, W. Bin, C. Wei, M. Sun, and Q. Jin. 2006. The use of comparative genomic hybridization to characterize genome dynamics and diversity among the serotypes of *Shigella*. BMC Genomics **7:**218.

Peng, J., Yang, J., Jin, Q. 2009. The molecular evolutionary history of *Shigella* spp. and enteroinvasive *Escherichia coli*. Infect. Genet. Evol. **9**, 147e152.

Prosseda G., M. L. Di Martino, R. Campilongo, R. Fioravanti, G. Micheli, M. Casalino, B. Colonna. 2012. Shedding of genes that interfere with the pathogenic lifestyle: the *Shigella* model. Research in Microbiology http://dx.doi.org/10.1016/j.resmic.2012.07.004.

Prosseda, G., Latella, M.C., Barbagallo, M., Nicoletti, M., Al Kassas, R., Casalino, M., Colonna, B. 2007. The two-faced role of cad genes in the virulence of pathogenic Escherichia coli. Res. Microbiol. **158**, 487-493

Prunier, A. L., R. Schuch, R. E. Fernandez, K. L. Mumy, H. Kohler, B. A. McCormick, and A. T. Maurelli. 2007a. *nadA* and *nadB* of *Shigella flexneri* 5a are antivirulence loci responsible for the synthesis of quinolinate, a small molecule inhibitor of *Shigella* pathogenicity. Microbiology **153:**2363–2372.

Prunier, A.-L., R. Schuch, R. E. Ferna´ndez, and A. T. Maurelli. 2007b. Genetic structure of the *nadA* and *nadB* antivirulence loci in *Shigella* spp. J. Bacteriol. **189:**6482–6486.

Pupo GM, Lan R, Reeves PR. 2000. Multiple independent origins of *Shigella* clones of *Escherichia coli* andconvergent evolution of many of their characteristics. Proc Natl Acad Sci U S A. 12;97(19):10567-72.

Pupo GM, Karaolis DK, Lan R & Reeves PR. 1997. Evolutionary relationships among pathogenic and nonpathogenic *Escherichia coli* strains inferred from multilocus enzyme electrophoresis and mdh sequence studies. Infect Immun **65**: 2685–2692.

Purdy, G.E., Fisher, C.R., Payne, S.M. 2007. IcsA surface presentation in *Shigella flexneri* requires the periplasmic chaperones DegP, Skp, and SurA. J. Bacteriol. 189, 5566-5573.

Ray, K., Marteyn, B., Sansonetti, P.J., Tang, C.M. 2009. Life on the inside: the intracellular lifestyle of cytosolic bacteria. Nat. Rev. Microbiol. **7**, 333-340.

Rowe, B., R. J. Gross, and D. P. Woodroof. 1977. Proposal to recognize serovar 145/146 (synonyms: 147, *Shigella* 13, *Shigella sofia*, and *Shigella manolovii*) as a new *Escherichia coli* O group, O164. Int. J. Syst. Bacteriol. **27:**15–18.

Sakellaris,H., Hannink,N.K., Rajakumar,K., Bulach,D., Hunt,M., Sasakawa,C., and Adler,B. 2000. Curli loci of *Shigella* spp. Infect. Immun. 68, 3780-3783.

Sansonetti, P. J. 2001. Rupture, invasion and inflammatory destruction of the intestinal barrier by *Shigella*, making sense of prokaryote-eukaryote cross-talks. FEMS Microbiol. Rev. **25:**3–14.

Sansonetti, P. J. 2006. Shigellosis: an old disease in new clothes? PLoSMed. **3:**e354.

Sansonetti, P. J., D. J. Kopecko, and S. B. Formal. 1982. Involvement of a plasmid in the invasive ability of *Shigella flexneri*. Infect. Immun. **35:**852–860.

Sasakawa, C., K. Kamata, T. Sakai, S. Makino, M. Yamada, N. Okada, and M. Yoshikawa. 1988. Virulence-associated genetic regions comprising 31 kilobases of the 230-kilobase plasmid in *Shigella flexneri* 2a. J. Bacteriol. **70:**2480–2484.

Schroeder G. N. and Hilbi H., 2008. Molecular patogenesis of *Shigella* spp.: controlling host cell signalling, invasion, and death by type III secretion. Institute of Microbiology p 134-156.

Schmidt H & Hensel M. 2004. Pathogenicity islands in bacterial pathogenesis. Clin Microbiol Rev **17**: 14–56.

Shepherd, J. G., L. Wang, and P. R. Reeves. 2000. Comparison of O-antigen gene clusters of *Escherichia coli* (*Shigella*) Sonnei and *Plesiomonas shigelloides* O17: Sonnei gained its current plasmid-borne O-antigen genes from *P. shigelloides* in a recent event. Infect. Immun. **68:**6056–6061.

Torres, A.G. 2008. The cad locus: more than just lysine decarboxylation. Anaerobe 15, 1-6.

Tritz, G. J., Chandler J. LR. 1972. Recognition of a gene involved in the regulation of nicotinamide adenine dinucleotide biosynresis. Journ. Of Bacteriology. **114:**128-136

Turner, S. A., S. N. Luck, H. Sakellaris, K. Rajakumar, and B. Adler. 2003. Molecular epidemiology of the SRL pathogenicity island. Antimicrob. Agents Chemother. 47:727–734. 314.

Turner, S. A., S. N. Luck, H. Sakellaris, K. Rajakumar, and B. Adler. 2001. Nested deletions of the SRL pathogenicity island of Shigella flexneri 2a. J. Bacteriol. 183:5535–5543.

Venkatesan, M. M., J. M. Buysse, and E. V. Oaks. 1992. Surface presentation of *Shigella flexneri* invasion plasmid antigens requires the products of the *spa* locus. J. Bacteriol. **174:**1990–2001.

Venkatesan, M. M., M. B. Goldberg, D. J. Rose, E. J. Grotbeck, V. Burland, and F. R. Blattner. 2001. Complete DNA sequence and analysis of the large virulence plasmid of *Shigella flexneri*. Infect. Immun. **69:**3271–3285.

Yang, F., J. Yang, X. Zhang, L. Chen, Y. Jiang, Y. Yan, X. Tang, J. Wang, Z. Xiong, J. Dong, Y. Xue, Y. Zhu, X. Xu, L. Sun, S. Chen, H. Nie, J. Peng, J. Xu, Y. Wang, Z. Yuan, Y. Wen, Z. Yao, Y. Shen, B. Qiang, Y. Hou, J. Yu, and Q. Jin. 2005. Genome dynamics and diversity of *Shigella* species, the etiologic agents of bacillary dysentery. Nucleic Acids Res. **33:**6445–6458

Yang, J., H. Nie, L. Chen, X. Zhang, F. Yang, X. Xu, Y. Zhu, J. Yu, and Q. Jin. 2007. Revisiting the molecular evolutionary history of *Shigella* spp. J. Mol. Evol. **64:**71–79.

Zhao, G., Zhu, L., Feng, E., Cao, X., Shang, N., Liu, X., Liao, X., Ying, T., Wang, J., Chen, H., Wang, H. 2010**.** A novel anti-virulence gene revealed by proteomic analysis in *Shigella flexneri* 2a. Proteome Sci. 8, 30.