1 Introduction

1.1 General Introduction

The Enterobacteriaceae form a large family of rod-shaped, facultative anaerobic, Gram-negative bacteria that are able to ferment sugars to produce lactic acid. Members of the Enterobacteriaceae are found in water, soil and in the gut flora of humans and animals. They include several renowned pathogens such as *Salmonella* and *Shigella* as well as opportunistic commensals such as *Escherichia coli*. Infections associated with the Enterobacteriaceae can be caused by endogenous bacteria that are normally part of the intestinal flora (commensals or opportunistic pathogens) or by exogenous bacteria derived from the environment. In many cases enterobacteria are associated with nosocomial infections that are acquired in hospital. For example, pneumonia, ear, sinus and urinary tract infections can be caused by *Klebsiella* and *Proteus* whereas bacteraemia can be caused by *Enterobacter* and *Serratia*. Enterobacteria-associated diarrhoea is a common problem and can range from minor nuisance to a life-threatening disorder especially in infants, the elderly, HIV infected patients and malnourished persons. The economic and health burdens exerted by enteric disease is enormous, although the exact amount is difficult to estimate (WHO, 2002). Entreobacteriaceae are found worldwide and are one of the leading killers of children in developing countries (Frey, 2002). For this reason it is important to understand the mechanisms of infection employed by these pathogens in order to design better control methods and therapies.

The enteric bacteria have evolved subtle mechanisms for interacting with the tissues of their hosts. For example, they can adhere to or invade eukaryotic cells via specific receptors found on the target cells. Some enteric bacteria harbour novel gene loci that express specific virulence associated factors that can be utilised during infection. Toxins can be produced to disrupt normal cellular processes. Consequently pathogenic enterobacteria can, ironically, be used as powerful probes to explore eukaryotic cell function. Historically, cellular studies on pathogenic bacteria have focused on the use of terminally differentiated cells adapted to long term survival in tissue culture. Some work has also focused on primary cells obtained directly from the host. Both these approaches have proved to be extremely productive and informative. However recent

work on pluripotent stem cells has shown their versatility as tools to explore biology but little has been done with these cells in terms of exploring pathogen interactions.

One of the aims of this project was to initiate studies on interactions between pathogens and murine stem cells with a view to gauging their potential as research tools.

1.2 Host-pathogen interaction

1.2.1 Infectious diseases today

The presence of infectious disease in our society is a sentinel of civilization (Gershon, 2000) and humanity has made great progress in the knowledge and defeat of transmitted disease. However in the world today, with an increasing number of people able to move easily from one continent to another together with the ability of the pioneers to reach unexplored ecosystems, society is facing increasing challenges from fast-tracking diseases and new potential pathogens. Technological progress has not completely eliminated the risk of disease transmitted from live stock animals which are nowadays often exchanged between countries and the emergence of zoonotic disease is still a front row factor in today's battle against infectious disease. For example in 1989 a new influenza strain (H3N2) was isolated in China and within 16 months was reported in Europe and North America (Meslin *et al.*, 1998). Moreover new technology can bring potential new sources of infection, for example xeno-transplantation practice could induce transmission of infectious agents originating in animals to human recipients (Domenech *et al.*, 2006). It may be possible to find a common denominator to these problems: globalization, a process bringing challenges to both industrialized and developing countries. Currently sanitary and phyto-sanitary control measures are a high priority in intercontinental travelling as well as in international trade. However it is not possible to reach a global high standard of precautionary measures, especially in developing countries where populations are still facing high rates of death from diseases commonly cured in industrialized countries. In fact more than 90% of the deaths from infectious diseases are caused by a few widely spread infections: lower respiratory infections, human immunodeficiency virus (HIV), diarrhoeal diseases, tuberculosis, malaria and measles (Coloma & Harris, 2008). In 2008 malaria, measles, lower respiratory infections and diarrhoeal illness are still causes of high mortality in infants and children in developing countries (Kaler, 2008). Malaria is an anopheline mosquito-borne parasitic disease; the most acute form is due to *Plasmodium falciparum.* The common symptoms of this disease are high fever, chill, vomiting and severe headache. A global estimation of 300 to 500 million clinical cases per year with more than 1 million paediatric deaths per year makes this disease one of the principal health

concerns. Measles is caused by a negative-strand RNA virus, specifically a *paramyxovirus* of the genus *morbillivirus*. The virus is transmitted by aerosol through the respiratory tract and induces high fever, cough, runny nose, conjunctivitis and characteristic skin rash. Measles represents one of the principal causes of vaccine-preventable childhood mortality and after a vaccination campaign measles deaths in Africa dropped by 60% in 2005 from 873,000 to 345,000 (UNICEF, 2005). Pneumonia is an acute lower respiratory tract infection caused by bacteria such as *Streptococcus pneumoniae*, several viruses or even fungi (Nascimento-Carvalho *et al.*, 2008). In children the majority of deaths occur in the newborn with about 1.6 million deaths in infants and children per year, of which 90% occur in developing countries. Many of the deaths caused by *S. pneumoniae* are 'vaccine preventable' and a clinical trial using a conjugated vaccine against 13 serotypes demonstrated an efficacy of between 50-80% of all paediatric deaths (Kaler, 2008). Diarrhoeal infections are caused by ingestion of virus, bacteria or parasites present in the water or food. Infection can spread easily into the population through contaminated utensils, hands or vectors such as flies. About 1.8 million children younger than 5 years die of diarrhoeal infections annually. Here, vaccination can have an impact but these are not available for some of the more common forms of diarrhoeal infection. Consequently, vaccines cannot substitute for the preventive practice of providing drinkable water to the population and following proper hygiene procedures (Kaler, 2008).

The incidence of these high burden diseases is enhanced by malnutrition, particularly associated with anaemia, and the spread of HIV (Pasquali, 2004). In Africa, the years of HIV pandemic have been superimposed on an enduring malaria pandemic. In Eastern and Southern Africa nearly 30% of the population is affected by the HIV virus and it is estimated that almost a quarter of clinical malaria occurs in HIV infected adults. HIV infection increases the incidence and severity of clinical malaria, however malaria seems to slow the effects of HIV. The enhancement and the interference of different combinations of infection is still under investigation and requires more emphasis (UNICEF, 2003). In developed countries many infectious diseases have been controlled or nearly eliminated as the main cause of mortality by the use of vaccines. Although vaccines have proved to be efficacious in the battle to control infectious disease the vaccine sector represents less that 2% of the world wide pharmaceutical market

(Editorial, 2008). Less than 10% of global applied research funding is directed to investigating the diseases that affect 90% of the population (Abbasi, 2001).

Projections for 2002 to 2030 predict that the risk of death in children younger than 5 years will fall by only 50% and global deaths due to HIV will rise from 2.8 million in 2002 to 6.5 million in 2030 (assuming that antiviral drug coverage will reach 80% by 2012). In an optimistic scenario, which takes into consideration improved preventive measures, HIV deaths are projected to increase to 3.7 million in 2030 (Mathers & Loncar, 2006). Taken together, these sorts of statistics highlight the requirement for more focused and effective research in this area.

1.2.2 Methods to study host-pathogen interactions

1.2.2.1 Historical

Koch's postulates are famous as a pioneering vision into infectious diseases. In 1890 Koch had already recognized the importance of the host as a key part of the pathogen survival, reproduction and persistence (Walker *et al.*, 2006). After all, there is no pathogen without a host. One of the first attempts to develop *in vitro* techniques to examine host/pathogen interactions was made in 1902 by the German botanist Haberlandt. He advanced the idea of cultivating artificial embryos from vegetative cells and demonstrated the toti-potency of plant cells to produce a whole new plant. Haberlandt reported for the first time the *in vitro* culture of isolated pallisade cells, pith cells, stamen hairs and stromal guard cells in enriched medium containing glucose (Werbrouck *et al.*, 1998). However in his efforts to demonstrate toti-potency the cells grown under these conditions did not divide but they did survive for a few weeks. In 1934, White was the first to obtain indefinite cultures with plant roots but Bonner in 1937 discovered that yeast extract components were fundamental to cell survival (Werbrouck *et al.*, 1998). Pioneering work on animal cells was done by Walther Flemming in the late 19th century and he initiated the science of cytology (Singer, 1989). The major contribution of *in vitro* culture of animal cells to infectious diseases has been to enable the growth of infectious agents that can't grow on agar or broth (Buehring, 1996). Animal cell culture, now referred to as tissue culture, became a routine laboratory technique in the 1950s. Major epidemics of polio in the 1940s and

1950s promoted efforts to develop effective vaccines and the need for cell culture became evident with the search for viral vaccines. The polio vaccine produced from deactivated virus became one of the first commercial products from cultured animal cells (Chaudry, 2004).

1.2.2.2 The use of genetics

The use of genetically modified organisms is common nowadays in many fields of research. Genome manipulation can help to characterize the role of genes in the ability of a micro-organism to infect and cause disease. On the other hand the manipulation of the host genome can indicate if a gene is involved in susceptibility or resistance to infections. In 1988 Professor Falkow proposed a revised version of the Koch's postulates adapted to molecular genetic studies on microbial pathogenesis that could help to identify genes related to pathogenicity. He proposed that genetic manipulation was almost a prerequisite for success and called this Molecular Koch's Postulates (Falkow, 1988). Molecular Koch's Postulates can be even more true for the host cell. Nevertheless, there is no doubt that this postulate contributed to the design of work leading to the identification of virulence factors that enable a bacteria to be pathogenic, or host 'factors' that permit a microbe to be virulent.

With the advance of research in this field it became more obvious that genetic manipulation has to be applied to both parties in infection and only in this way is it possible to obtain a complete picture of the complex interactions between host and pathogen. For example, the SipB protein encoded on the Salmonella Pathogenicity Island 1 (SPI-1) and secreted via a Type III secretion system (TIIISS) in *Salmonella* was suspected to interact with the host protein Caspase-1. When the gene encoding this protein was mutated, the bacteria were unable to express a functional SPI-1 and could not be used to confirm any interaction with Caspase-1. In this case the use of caspase-1 knockout mice, together with biochemical interaction assays, helped to define the role of SipB and the relationships with Caspase (Jarvelainen *et al.*, 2003; Lara-Tejero *et al.*, 2006; Monack *et al.*, 2000). Macrophages were subsequently harvested from mutant mice and these were used to further explore the biology of these interactions.

Mutagenesis of the host genome is now really in vogue and new mutagenesis techniques have been developed to investigate this complex system. Lengeling *et al.* recognize

three main approaches that are already in place: 1) target mutations of candidate loci to disrupt specific immune-responses; 2) identification of new loci involved in host response through induced or spontaneous mutation; 3) induction of new inbred mouse strains that are susceptible or resistant to specific pathogens, this will permit the tracing of complex gene interactions using quantitative trait loci (QTL) mapping (Lengeling *et al.*, 2001). The combination of host and pathogen genetic manipulation has been referred to as "genetics-squared" (Persson & Vance, 2007).

1.2.2.3 The use of cell culture

Models have played an important role in the study of infectious disease particularly as it is not ethical to expose humans to potentially lethal pathogens in order to study disease progression (Wiles *et al.*, 2006). *In vivo* animal models have also contributed in a great way to advances in infectious disease research although there are also ethical issues here. Consequently, the use of *in vitro* models has proved to be of great value in this field. Both *in vivo* and *in vitro* models cover different aspects of host-pathogen interactions and they complement each other. On one hand animal models can give insight into the complex interactions that a biological system exhibits during infection, including synergic responses between different components of the body. On the other hand *in vitro* models offer the possibility of investigating host-pathogen interactions at the cellular level and examine the cellular reaction to pathogen invasion. Both of these approaches have pros and cons and they cannot fully replace the natural host. *In vitro* models are advantageous in many ways, they can be well controlled, they are comparatively inexpensive and they have limited ethical issues. Further, human cells can be employed. The examination of host-pathogen interaction *in vitro* can be conducted under well controlled conditions and close observation is possible. The use of live animals in experimentation is associated with ethical responsibilities and substantial administrative costs. Many countries participate in the three Rs: replacement, refinement and reduction. However, it is not always possible to eliminate *in vivo* models.

In vitro models can be classified into three major types: non-polarized cell layers, polarized cell layers and 'organoids'. The immortalised properties of non-polarized cell models raise some concern about their 'normal' behaviour during infection. Polarized

cells are regarded as being generally more closely related to the *in vivo* situation, however aberrant genomic composition and properties can effect the outcomes of host-pathogen interactions. The third class of *in vitro* model 'organoids' represent the only three dimensional *in vitro* model and are perhaps the closest to *in vivo* structures (Wiles *et al.*, 2006). ES cells may represent a revolutionary *in vitro* model that can avoid some of the pitfalls of existing *in vitro* cell culture models. One important characteristic of ES cells is that they have and retain a normal karyotype. Further, human ES cells can be used to investigate species-specific pathogens that cannot be effectively used in other models. Additionally, the pluripotency and flexibility of these cells can provide the opportunity to develop three dimensional *in vitro* models that might reproduce more closely *ex-vivo* structures.

If ES cells are to be exploited as novel cells for investigating host-pathogen interactions it is worth considering which pathogens might be exploited in such systems. Some pathogens such as *Salmonella* and *Shigella* can actively invade human and murine cells using specialised genetic systems encoded in their genomes. Also, unlike viruses, intracellular growth is not intimately linked to viability as bacteria can readily survive outside of cells. Hence *Shigella* and *Salmonella* can be considered as potentially ideal probes for investigating pathogen interactions with ES cells. Mutant bacteria, as well as mutant ES cells, are available and the interaction of *Shigella* and *Salmonella* with differentiated cells has been well described. Consequently, it is worth looking here in some detail at the pathogenic properties of these bacteria.

1.3 *Shigella*

1.3.1 Bacteria and disease

Shigella, members of the Enterobacteriaceae, are uncapsulated, non-motile bacteria responsible for a significant global human health problem. *Shigella* can be classified in four species *S. dysenteriae*, *S. flexneri*, *S. boydii* and *S. sonnei*, these species are further subdivided into serotypes on the basic of the surface O-specific polysaccharide LPS. *Shigella dysenteriae* type 1 was the first Shigella species to be isolated in 1896 by Kiyoshi Shiga (Niyogi, 2005). *Shigella* genomes are very similar to *Escherichia coli* although comparatively they have significant chromosomal rearrangements and unique genetic islands called *Shigella*-specific islands (Sis). Many Sis have the classical characteristics of pathogenicity islands. In addition *Shigella* can harbour different plasmids, some of which are associated with invasiveness. *Shigella* can be found in the environment and they can survive for weeks in cool and humid locations or for up to 46 days when dried on linen (Altwegg & Bockemuhl, 1998). Moreover, *Shigella* survive and divide in humans, the only natural host, causing shigellosis or dysentery. They have a low infectious dose and consequently they can easily spread from person to person and be transmitted via the environment. The principal form of transmission is by the faecal-oral route. Shigellosis symptoms include dysentery with frequent mucoid bloody stools, abdominal cramps and tenesmus. It has been estimated that there are 163 million episodes of shigellosis per year with more than one million deaths (Kweon, 2008). No definite groups of individuals are immune to shigellosis but certain individuals are at higher risk. Most episodes of shigellosis in healthy individuals are self-limited and resolve within 5-7 days without consequences. The incidence of shigellosis is elevated among children of 1 to 4 years old and in those that suffer from malnutrition the disease can be quite harsh and induce further growth retardation. *S. sonnei* and *S. flexneri* are linked to the endemic form of the disease which occurs worldwide. *S. flexneri* is the hyper-endemic type in developing countries and is accountable for 10% of all diarrhoeal episodes among children younger than five years (Niyogi, 2005). *S. dysenteriae* 1, which expresses the potent Shiga toxin, is responsible for the epidemic form of the disease which accounts for deadly outbreaks in developing countries. Shigellosis diagnosis requires laboratory growth of stool sample on differential/selective media and

aerobic incubation to inhibit the growth of anaerobic normal flora. This clearly makes diagnosis quite a challenge in developing countries. Vaccines effective against *Shigella* are not available despite the international community efforts in this direction. There is need for a multivalent vaccine covering the prevalent species and serotypes. One of the major obstacles to vaccine development is the unsuitability of animal models. For a comprehensive review refer to Kweon (Kweon, 2008).

1.3.2 Mechanism of host invasion

1.3.2.1 Type III secretion

Shigella are able to invade and multiply within colonic epithelial cells and the rectal mucosa, causing cell death while spreading laterally through the epithelial layers (Niyogi, 2005). Before the development of dysentery symptoms include early inflammatory lesions of the colorectal mucosa, comparable to aphthoid ulcers, with the presence of lymphoid follicles (Sansonetti, 2002). *Shigella* predominantly invades the colonic epithelium through M cells in the Peyer's patches, which do not bear a microvillae brush border (Jepson & Clark, 2001). Also, *Shigella* can be phagocytosed by dendritic cells (DC) and macrophages resident in the dome. Subsequently, intracellular *Shigella* are able to induce macrophage apoptosis involving caspase-1 activation and consequent secretion of two cytokines, IL-1β and IL-18 (Sansonetti, 2002).

Shigella species. harbour plasmids that confer the 'invasive phenotype'. The plasmid pWR100 (214kb) from *S. flexneri* serotype 5a strain M90T has been fully sequenced and it contains a pathogenicity island of 30kb encoding the *ipa/mxi-spa* operons which direct the expression of Ipa or invasin proteins delivered inside the host cell through the TIIISS apparatus formed by the *mxi* and *spa* products (Buchrieser *et al.*, 2000). The primary function of the TIIISS is to transport proteins from the bacterial cytoplasm into the host cell. The TIIISS molecular structure was described as a 'needle complex' (NC) by Kubori and coworkers in 1998 (Kubori *et al.*, 1998) and even though *Shigella* is a non-motile bacterium the structure resembles flagellar basal bodies (Blocker *et al.*, 2001). Three parts of this secretion system were described: an external needle, a transmembrane complex or neck and a large bulb-like structure internal to the bacterial membrane. The needle complex is formed by at least five proteins: MxiD, MxiG, MxiJ,

MxiH, and MxiI (Blocker *et al.*, 2001). The secreted proteins remain in the bacterial cytoplasm until the secretion machinery is activated by host contact or external signals such as serum or the small amphipathic dye molecule Congo red (Bahrani *et al.*, 1997; Menard *et al.*, 1994). Then at least three invasion proteins are injected inside the host cells (IpaB, IpaC and IpaD) through a channel within the structure (Blocker *et al.*, 1999). The Ipas are able to catalyze the formation of a localized actin rich macropinocytic-like ruffle of the host cell membrane which mediates bacterial internalization (Bourdet-Sicard *et al.*, 1999; Niebuhr *et al.*, 2000; Tran Van Nhieu *et al.*, 1999). Once inside the cell the *Shigella* bacterium is able to escape the vacuole and start a cycle of intra- and inter-cellular spreading.

1.3.2.2 Intracellular lifestyle

Shigella is an invasive bacterium able to induce its own phagocytosis and survive and spread from cell to cell disrupting the colonic epithelial structure and inducing inflammation and dysentery. It has been established that the *Shigella* internal-cell growth rate is not dependent on the Shiga toxin but on the plasmid-mediated dynamic lysis of the phagocytic membrane soon after endocytosis (Sansonetti *et al.*, 1986). The endosomal membrane lysis is mediated by Ipa proteins injected into the host cell (High *et al.*, 1992). IpaB seems to play the main role and a *Shigella ipaB* mutant showed cell entry but no lysis of the phagocytic vacuole. This *Shigella* mutant did not induce the contact mediated haemolytic activity characteristic of this bacterium. Once free inside the host cytoplasm the bacteria move via the formation of an F-actin tail. The bacterial movements are random and quick and the catalyst for this movement is the polymerization of actin at the base of the bacteria. The actin polymerization is due to the IcsA protein which is encoded on the large virulent plasmid pWR100 (Bernardini *et al.*, 1989). IcsA distribution on the bacteria surface is asymmetric and accumulates at the bacterial pole. This polar localization is also accentuated by the action of a protease, SopA, which seems to regulate the asymmetric distribution of IcsA. *Shigella sopA* mutants do not shown intracellular motility (Egile *et al.*, 1997). The random propulsion of the bacteria inside the host cell cytosol ultimately causes a contact with the inner face of the cytoplasmic membrane and the bacteria protrude into the adjacent cell until they are eventually phagocytosed. It has been speculated that the bacteria establish contact with components of the cellular junction and that the passage is an active actin-driven

protrusion. However, the interaction with cellular cadherins is a prerequisite to inter-cellular spreading of *Shigella* (Sansonetti *et al.*, 1994). Once engulfed in the neighbouring cell, *Shigella* is surrounded by a double membrane vesicle and in this case liberation into the cytosol is due to IcsB, which is encoded just up-stream of the Ipa genes in the locus encoding the effectors for the entry (Allaoui *et al.*, 1992).

All this happens in epithelial cells. *Shigella* follows a different behaviour once inside macrophage cells located just underneath the Peyer's patches. Once inside macrophage cells, *Shigella* induces host-cell apoptosis through the action of IpaB protein which activates caspase-1 (Hilbi *et al.*, 1998). Caspase-1 activation mediates the cleavage of the pro-inflammatory cytokines, IL-1β and IL-18, triggering the inflammation reaction probably through the production of IFNγ. Once outside the host-cell the bacteria are more vulnerable to host defences. Neutrophils can produce elastase which actively degrades virulence factors and prevents the escape of *Shigella* from the phagocyte (Weinrauch *et al.*, 2002). Also the intracellular bacteria can be trapped in intertwined filamentous structures known as neutrophil extracellular traps preventing infection of intact cells (Brinkmann *et al.*, 2004). In this condition the bacteria can be kept under control and the host immune system can take action against the pathogen.

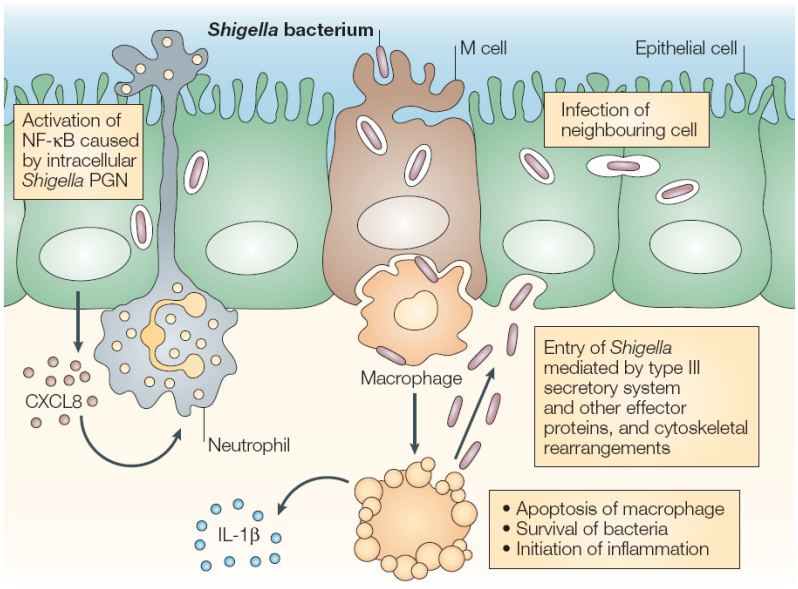

Figure 1.1 ***Shigella* invasion of the gut mucosa**

Shigella invades the gut epithelium preferentially targeting M-cells associated with the follicle-associated epithelium above the lymph node. Bacteria can spread laterally invading epithelial cells or phagocytic

cells present beneath the follicle associated epithelium (FAE). In macrophages apoptosis can be induced through the activation of caspase 1 and the release of mature IL-1β and IL-18 and CXC-chemokine ligand 8 can then initiate the inflammation response. (Sansonetti, 2004)

1.4 *Salmonella*

1.4.1 History

Salmonella was named after Daniel Elmer Salmon, an American veterinary pathologist who first discovered the *Salmonella* bacterium in pigs in 1885. However typhoid fever was first described by the French physician Pierre Charles Alexandre Louis who first proposed the name "typhoid fever" as early as 1827. Interestingly, it is believed that several historical figures were killed by Salmonellosis, including Alexander the Great. He mysteriously died in 323 B.C. and in 1998 a group from the University of Maryland School of Medicine reported that his death was likely due to typhoid fever (Oldach *et al.*, 1998). In recent years many cases of Salmonellosis have been reported throughout the western world, where *Salmonella* outbreaks are associated not only with eggs and cow milk but with canned products like peanut butter and tomatoes (http://www.salmonellablog.com/articles/salmonella-outbreaks/).

For example consumeraffairs.com reported three outbreaks of Salmonellosis in February 2007 in the US, Canada and UK respectively. It is difficult to generate a true global picture of outbreaks in developing countries where *Salmonella* overlaps with other diseases but the problem is significant.

1.4.2 Bacteria and disease

The *Salmonella* genus is divided in two species: *Salmonella bongori* and *Salmonella enterica* which is further subdivided into five subspecies: I - enterica, II – salamae, IIIa – arizonae, IIIb – diarizonae, IV – houtenae, V – obsolete , VI – indica. *Salmonella* are today classified according to the serological techniques developed by Kauffman and White, which uses antisera to determine the taxonomical characteristic of the surface O antigen which is the polysaccharide associated with the lipopolysaccharide of the bacterial outer membrane. Antisera to the flagellar antigen H is also used and this can be further divided into phase 1 and phase 2. These phase antigens reflect the ability of *Salmonella* to live in different phases: a motile phase and a non-motile phase or flagellated and non-flagellated phases (sometimes additional flagella are also expressed). *Salmonella* serovar Typhi expresses an additional polysaccharide antigen,

Vi, associated with the bacterial capsule. When expressed this antigen causes enhanced virulence as determined by a lower infective dose in human volunteers (Parry, 2006).

Salmonella Typhimurium has been recognised as a common cause of disease throughout the world and was first isolated in late 19th century (Forsyth, 1998). *S. enterica* serovar Typhimurium (*S.* Typhimurium) is an etiological cause of salmonellosis in humans and animals and is adapted to a life-style in the gastro-intestinal tract. *S.* Typhimurium is normally contracted by ingestion of contaminated food of animal origin like eggs and milk or from water. For this reason it is more commonly found in developing countries where water supply systems can be cross-contaminated by animal and human waste. This pathogen is able to escape the hostile stomach environment to attack the intestinal epithelium normally causing localised gastroenteritis in humans and systemic typhoid fever in mice. However, strains of *S.* Typhimurium (Non-typhoidal salmonellae or NTS) can cause invasive disease in humans in some regions of the world, particularly Africa. The symptoms of *S.* Typhimurium appear 12-72 hours post infection and include fever, abdominal pain, diarrhoea, nausea and sometimes vomiting. The illness normally lasts 4-7 days and most people recover without treatment, but in very young and elderly people the bacteria can enter the blood stream and antibiotic therapy may be needed (WHO, Fact Sheet N 139 Revised April 2005). In developing countries as in Africa, salmonellosis is more common during the rainy season but also can reflect seasonal patterns of diseases such as malnutrition and malaria. The incidence of sepsis caused by NTS is very high in African children. In 1987 a study was conducted on NTS bacteraemia in Rwanda where 72% of children with NTS bacteraemia were sick enough to require hospitalization and a 12% fatality rate was reported. Also, cases of NTS meningitis were reported in Malawi between 1996-1997 with a fatality rate of 57% (Graham *et al.*, 2000).

1.4.3 Type Three Secretion Systems

S. enterica has received the most attention since it is responsible for most human disease and contains many serovars associated with disease. *S. enterica* serovar Typhi is a human restricted serovar associated with systemic typhoid disease. *S.* Typhimurium causes systemic disease in the mouse and is often used as a surrogate model of typhoid fever. These two serovars exhibit significant genome homology. However, *S.* Typhi has

a high number of pseudogenes, 204, compared to *S.* Typhimurium and these mutations may contribute to host restriction. Genome comparisons with other enteric bacteria facilitate the identification of genetic insertions and deletions. Many of the larger insertions encode for virulence determinants that are named *Salmonella* pathogenicity islands (SPIs). Some SPIs are shared between all or few *Salmonella* serovars but a few are restricted to particular serovars. The TIIISS encoded on SPI-1 has a significantly lower GC content than the chromosome and is flanked by IS-3-like elements (Sukhan, 2000). TIIISSs have been found in many animal and plant pathogens and the fact that both animals and plants can be invaded by pathogens carrying this type of secretion system highlights their versatility (Sukhan, 2000). The horizontal insertion of these elements has been supported by three observations: different GC content compared to the rest of the bacterial chromosome; flanking genes that are contiguous in closely related non pathogenic bacteria; transposons or bacteriophage sequence close to SPIs suggesting mechanisms of acquisition (Ohl & Miller, 2001).

S. enterica encode at least two TIIISSs associated with the invasion process, located in SPI-1 and 2 respectively. SPI-1 encodes a system required for entry into epithelial cells and at least 13 known effector proteins, several of which are involved in the host cytoskeleton remodelling inducing bacteria phagocytosis, are secreted directly into host cells (Pascale Cossart, Second Edition 2005). SPI-2 encodes genes essential for intracellular replication and this locus is necessary for the establishment of systemic infections. Different TIIISS share common features such as the absence of a cleavable amino-terminal signal sequence on the secreted proteins, the need for specific chaperones for the secretion of effector proteins and the requirement of an activation signal involving contact with the host cells. Contact is required for the full formation of the system and delivery of proteins into the cytosol of host cells. Sequence comparisons have highlighted similarities between components of the TIIISS and the proteins involved in their export (Sukhan, 2000). Also electron microscopy has revealed an elongated structure spanning the inner and the outer membranes called the needle complex for both these systems. The TIIISS encoded on SPI-1 can be distinguished into structural components including a basal structure that resembles the flagellar basal body and an outer membrane structure named the needle complex (Kubori *et al.*, 1998). In addition there are a number of cytoplasmic proteins that are thought to help the assembly of the secretion apparatus. At least five proteins have been described that

make up part of the inner cell structure. These are InvA, SpaP, SpaQ, SpaR and SpaS (Hensel, 2006). Homologs of these proteins are found to play key roles in the flagellar assembly systems in various bacteria (Ohnishi *et al.*, 1997). *Salmonella* InvA is homologous to MxiA of *Shigella* (Ginocchio & Galan, 1995). Whereas many different effector proteins such as SipA, SipB, SipC, SptP are encoded by genes located within SPI-1, SopA, SopB, SopD, SopE and SopE2 are encoded in loci outside the island (Mirold *et al.*, 2001). Also, other proteins, such as those involved in iron uptake, are encoded on the locus (Zhou *et al.*, 1999).

The TIIISS encoded on SPI-2 is essential for *Salmonella* to proliferate within the host cell. The structure of this TIIISS has not been fully determined but needle like structures were determined to be present on *Salmonella* inside the SCV (Chakravortty *et al.*, 2005). SPI-2 encodes several translocator proteins including SseB, SseC and SseD and maybe SpiC (Freeman *et al.*, 2002), and a few effector proteins SseF, SseG (Kuhle & Hensel, 2002). Others such as SifA, SifB, SspH1, SspH2, SlrP, SseI, SesJ, PipB, PipB2 and SopD2 are encoded outside the SPI-2 locus (Hensel, 2006).

1.4.4 Mechanism of host invasion

Salmonella TIIISSs are responsible for both invasion and bacterial survival in the host cells. *Salmonella* invasion of non-phagocytic cells is mediated by the TIIISS encoded on SPI-1 whose secreted effector proteins induce cytoskeletal rearrangements that lead to membrane ruffling facilitating bacterial internalization. *Salmonella* invade the gut epithelium preferentially through M-cells. These cells are generally targeted by pathogens since they have a reduced amount of mucus on their surface and present irregular microvilli. M-cells are located in regions of follicle-associated epithelium (FAE) which are organized mucosa-associated lymphoid tissues such as the intestinal Peyer's patches (Jepson & Clark, 2001). The first effectors proteins to be translocated inside the host cells are SopE, SopE2 and SptP which target the monomeric GTP-binding proteins of the Rho family. Members of the Rho family include cdc42, rac and rho and they play a central role in regulating the actin cytoskeleton. SopE activates these GTP-ases inducing a reorganization of the cytoskeleton F-actin filaments. These form ruffles by activation of JNK kinase which ultimately regulates transcription factors (Patel & Galan, 2006). Also the SPI-1 TIIISS induces transmembrane fluid and

electrolyte fluxes in addition to the synthesis of cytokines and prostaglandin mediator factors of inflammation. Once inside the host cell, the pH change inside the endosome induces the expression of the SPI-2 TIIISS which is involved in bacterial survival and persistence inside the host cell. One effector protein playing an important role is SifA, which is responsible for the maintenance of the integrity of the SCV during intracellular survival (Beuzon *et al.*, 2000). Another protein secreted by SPI-2 TIIISS is SpiC that appears to interfere with normal cellular trafficking (Uchiya *et al.*, 1999)

In phagocytic cells *Salmonella* can induce rapid apoptosis, stimulated in part by the delivery of cytosolic flagellin through the assembly of the IpaF inflammasome leading to activation of caspase-1 and finally resulting in the activation of IL-1β and IL-18 (Franchi *et al.*, 2006; Miao *et al.*, 2006). *Salmonella* associated phagocytic cells produce and release antimicrobial products such as reactive oxygen species (ROS) and reactive nitrogen species (RNS) inducing local inflammation (Abrahams & Hensel, 2006).

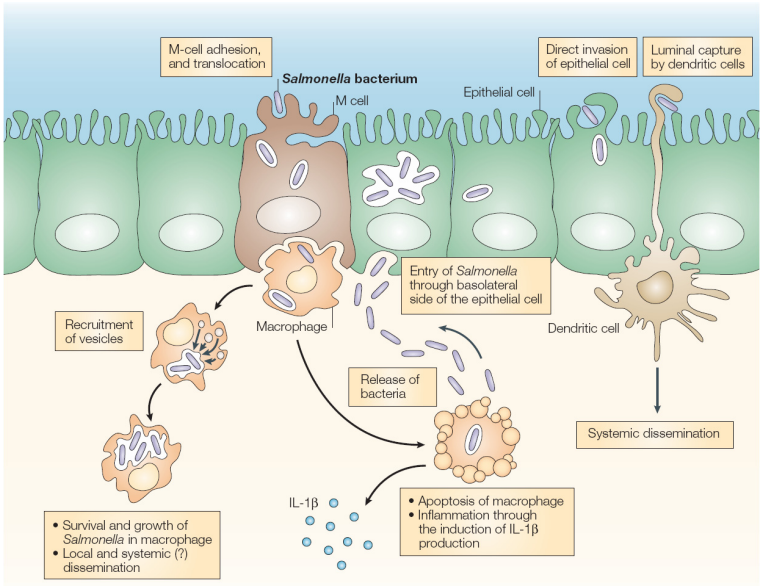

Figure 1.2 ***Salmonella*** **invasion of the gut mucosa**

Salmonella invade the gut epithelium through M-cells and reach the underlying lymph nodes where phagocytic cells take up the bacteria. Salmonella is able to induce apoptosis of macrophages and DCs with consequent secretion of IL-1β and IL-18. In this representation DCs are shown protruding through the epithelial cells to sample the lumen content (Sansonetti, 2004).

1.5 Embryonic stem cells

1.5.1 History

The first documented suggestion that cells had pluripotency goes back as far as 1902 when Haberlandt suspected that plant cells could produce a whole organism. However it proved to be more challenging to establish the same concept in animal cells due to the complicated pattern of functions and tissues found in the animal body. Nevertheless, the concept that cells can be pluripotent was not abandoned by researchers. The first reports of the formation of a mosaic mouse were in 1961 (Tarkowski, 1961) followed in 1981 by the description of the first establishment of an *in vitro* culture of pluripotent cells. Two laboratories obtained pluripotent cells isolated directly from the inner cell mass (ICM) of late blastocysts cultured in medium conditioned by established teratocarcinoma stem cell masses (Evans & Kaufman, 1981; Martin, 1981). The ES cultures were initially maintained in undifferentiated status by co-culture with mouse embryonic feeder cells. Later a factor produced by the feeder cells that contributed to ES cell growth was recognized to be the Leukaemia Inhibitory Factor (LIF) which is a chemokine of the IL-6 family that acts by activating STAT pathways (Smith *et al.*, 1988; Williams *et al.*, 1988). More recently other factors have been recognized as playing a pivotal role in maintaining the undifferentiated state of ES cells including the bone morphogenetic proteins (BMPs) and transcription factors such as Oct3/4 and nanog (Chambers & Smith, 2004). These discoveries were important in advancing the understanding of how ES cells maintain their self-renewal status and enabled researchers to grow ES cells in FCS free medium.

1.5.2 Murine ES cells

ES cells derive from the ICM of an embryo at the pre-implantation stage or blastocyst. The mouse embryo is in the blastula stage when it harbours between 50-250 cells or is at 3-4 days of growth. Two types of embryonic cells can be observed in the blastocyst, namely the ICM and the trophectoderm (TE). The TE is formed by large polar cells organized in an epithelium of single cells compactly connected by tight junctions, desmosomes and gap junctions whereas the ICM is formed by smaller, apolar cells

compacted to form the epiblast (Ducibella *et al.*, 1975; Ducibella & Anderson, 1975; Magnuson *et al.*, 1977). In later foetus development the TE descendant cells develop into extra-embryonic structures such as the placenta that makes contact with the mother body. The ICM grow into the three germinal layers: ectoderm, endoderm and mesoderm as well as into other external structures such as visceral yolk sac, amnion and the allantois that will form the umbilical cord (Lopez & Mummery, 2004). For this reason the cells that compose the blastocyst are identified as pluripotent because they can differentiate into multiple lineage cells.

ES cells have the characteristic of being self-renewing. This means that they are able to divide and create a copy of the original cell and progenitor cells are able to generate certain lineages of differentiated cells. In this latter case the differentiation capacity of the cell is restricted (Smith, 2001). Totipotent cells are typical of the embryo at the cleavage stage. As the embryo divides the blastomers lose the ability to differentiate into all cell types present in the foetus and in the adult. Three types of cell lines deriving from embryonic cells are currently available as cellular models. They are: Embryonal Carcinoma (EC) cells, Embryonic Stem (ES) cells and Embryonic Germ (EG) cells. All of these have pluripotent traits and a distinct ability to differentiate.

The term 'Stemness' refers to the transcriptional profile underlying the molecular process of the stem cell properties of self-renewal and the generation of differentiated progeny. Although stem cells in different cellular environments or niches will have different physiological demands and therefore distinct molecular programs, it is believed that a certain genetic profile is shared by all (Melton & Cowan, 2004). Some characteristic transduction pathways have been identified as active in stem cells (Clark *et al.*, 2004) and additional studies have been conducted in this direction (Ivanova *et al.*, 2002; Ramalho-Santos *et al.*, 2002)

1.5.3 Differentiating ES cells

The main feature of embryonic stem cells is the ability to differentiate into any kind of somatic cell present in the adult body. Their ability to 'produce' any type of cell is likely to find application in therapeutic treatment to cure organ failure or dysfunction. Thirty years after their initial *in vitro* culture, ES cells have been exploited in a vast

number of differentiation protocols. They can undergo spontaneous differentiation *in vitro*, by forming embryoid bodies (EBs) that contain different types of cells (Abe *et al.*, 1996) and *in vivo*, following their injection into immunodeficient mice, form differentiated tumors known as teratocarcinomas (Stevens, 1984). *In vitro* differentiation methods can be distinguished in three different approaches. In the first approach the ES cells are grown in single-cell-suspension and dividing, they form three-dimensional structures called embryoid bodies (EBs) (Doetschman *et al.*, 1985; Keller, 1995). In the second method the ES cells are grown on a stromal cell layer and the differentiation occurs in contact with these cells (Nakano *et al.*, 1994). The most widely used stromal cells for the last method are OP9 which were derived from a CSF-1 deficient op/op mouse (Yoshida *et al.*, 1990). The third protocol involves ES cells grown in a monolayer on extra-cellular matrix proteins (Nishikawa *et al.*, 1998). The three different protocols have pros and cons and not all of these methods are optimized for all types of differentiation processes that can be performed. The first two methods listed have been described as efficient protocols. However the third is the most appealing for therapeutic application since it can deliver pure populations of differentiated cells without any need for screening different type of cells. However, this methodology still needs to be optimized for most applications. One of the main factors that stand between the *in vitro* differentiation and the *in vivo* implantation is the use of FCS in the differentiation protocols. FCS has an uncertain composition and is subjected to composition variability from batch-to-batch. Also, human re-implantation of cells differentiated in FCS can present some immunological and ethical issues. Nevertheless the research on *in vitro* differentiation protocols has made huge progress in the past 10-15 years. Differentiated cells deriving from *in vitro* cultured ES cells can be distinguished based on the germinal layer they are derived from. Among the mesoderm-derived lineages that have been explored, protocols to obtain haematopoietic, vascular and cardiac cells appear to be the easiest to perform but skeletal muscle, osteogenic and adipogenic cell protocols have also been described. Research on lineages with endoderm origins is concentrated especially on pancreatic β-cells and hepatocytes, with the hope of their clinical applicability to treat type I diabetes and liver diseases. The ectoderm differentiation of mouse ES cells is well established and many studies have reported the characterization of neural differentiation. Although these are interesting subjects, a detailed discussion of the different differentiation protocols and clinical

applications goes beyond the scope of this study so for an extensive review on the subject please refer to Keller (Keller, 2005).

1.5.4 ES cell mutants and KO mice

Many researchers have documented the importance of ES cells and the many advantages they present for clinical applications as well as for basic scientific research. Indeed, perhaps the biggest advantages of ES cells versus immortalized cell lines is that they maintain a normal karyotype following extensive passaging in culture and have the propensity to be genetically manipulated (Austin *et al.*, 2004). The genetic manipulation of a specific gene in ES cells can be exploited to investigate function during development, if expression is followed during embryoid body formation or if the mutant cell is employed to generate chimeric mice. Also, the gene mutation can be investigated during lineage differentiation or in studies of cell homeostasis. Many of the methods that can be used to induce gene mutations in ES cells are summarized in Figure 1.3.

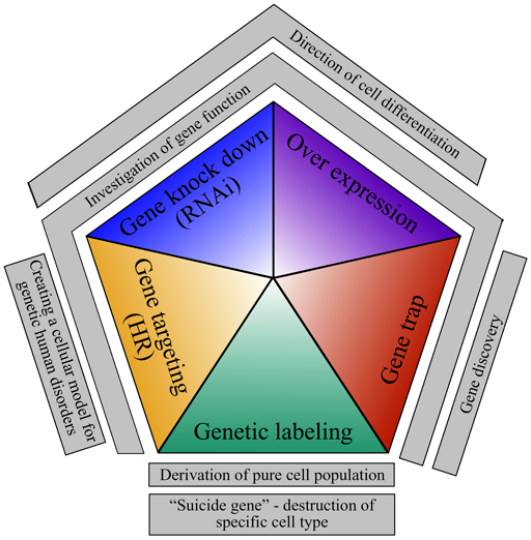

Figure 1.3 Genetic manipulation of ES cells and their potential application

This cartoon summarizes the genetic approaches that can be used to obtain mutant ES cells. The more traditional approaches, like transfection and the use of viral vectors can be more difficult to use especially on human ES cells. Transfection is the introduction of foreign DNA into the cultured cells by physical or biochemical methods (electroporation or LipofectAMINA Pulse, FuGENE). Infection using viral vectors, especially lentiviruses, have shown to provide a high level of stable integrants. However, the long terminal repeats (LTRs) may effect expression of the transgene, the vector size can be a limiting factor in addition to non-specific integration and potential effect of the virus genome in clinical studies. However, these methods still find use in mouse studies. Genetic tagging can be used to track transplanted cells or

the expression of a mutated gene. RNAi techniques use RNA to interfere with the gene expression through mRNA degradation. In this case researchers talk of knock-down expression by transfection of small RNA or of a plasmid that directs the production of interfering molecules (Kopper & Benvenisty, 2005). Gene trap is a random introduction of the mutagenic agent into the promoter, enhancer or the polyadenylation (poly-A) of a gene and is meant to disrupt the gene function. For an extensive review on gene-trapping please refer to Stanford, 2001 (Stanford *et al.*, 2001).

Of all the techniques reported in Figure 1.3 the main approach used to create ES cells mutants and mouse knockouts is Gene Targeting (Figure 1.4). The remarkable importance of this process has been underlined by the award of the 2007 Nobel Prize in physiology or medicine to three researchers, Mario R. Capecchi, Martin J. Evans and Oliver Smithies, for their discover of "principles for introducing specific gene modifications into mice by the use of embryonic stem cells" (The Nobel Prize Assembly, Press Release 2007-10-08). This was a milestone of modern research. In 1986 Capecchi and Smithies proved that it was possible to target specific genes by homologous recombination in cultured cells and Evans contributed by providing the necessary vehicle to target the mouse germ line to transmit the mutation from generation to generation, mouse ES cells. In 1989 the first knockout mice using gene targeting were produced. By 2004 about 10% of all mouse genes were targeted in the form of knockout mutant mice (Austin *et al.*, 2004).

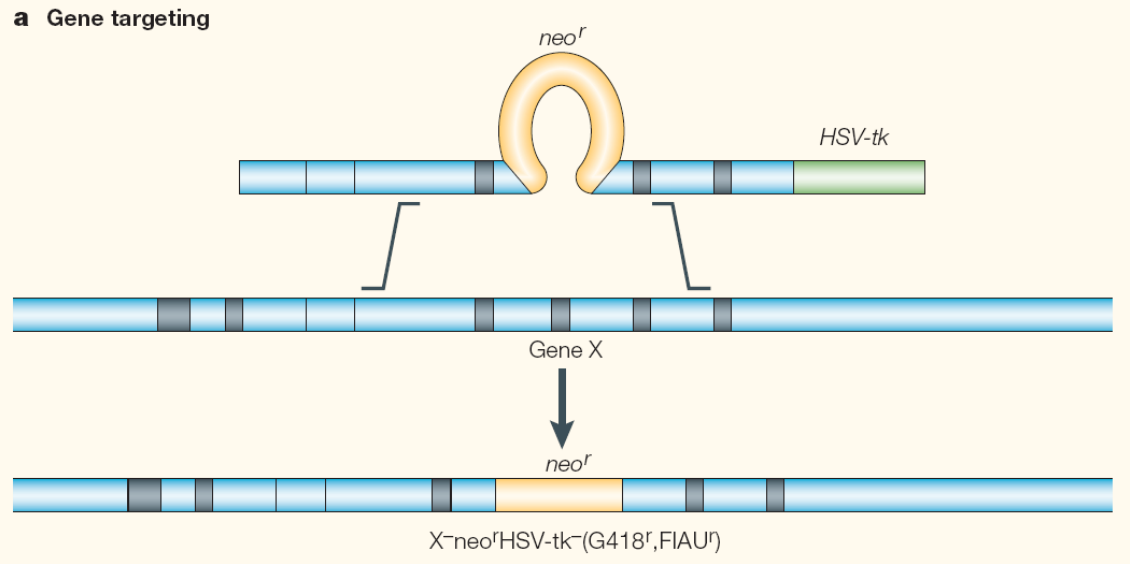

Figure 1.4 Generation of target mutant ES cells

The target vector is introduced by electroporation into ES cells and the desired mutant cells are selected by positive selection. The vector contains a neomycin resistance (neo^r) gene in an exon of the target gene and a linked herpes virus thymidine kinase (HSV-tk) gene at the other end. The vector pairs with the chromosomal target gene and after homologous recombination the resistance to neomicyn is inserted into the gene, disrupting one copy of it. The cells in which this event occurs will be of $X^{+/-}$, neo^{r+} and HSV-tk^- genotype and will be resistant both to G418 and FIAU (Capecchi, 2005).

Ambitious programmes for more efficient and effective production of mouse knockouts were announced in 2004 bringing together international efforts to achieve this ambitious and important goal for the research and social community (Auwerx *et al.*, 2004). In the last 10 years a new mutagenesis technique has been introduced that utilizes a bacterial artificial chromosome (BAC) to introduce mutagenesis into the mouse genome with more efficiency and less labour. Nowadays this technique is used to produce 'mutant libraries' that can be employed for the *in vivo* and *in vitro* study of specific genes in different mice backgrounds, for an example refer to Adams *et al.* (Adams *et al.*, 2005).

1.6 Dendritic Cells

1.6.1 What are they?

Dendritic cells (DCs) are professional antigen presenting cells that were described for the first time in 1973 by Steinman (Steinman & Cohn, 1973; Steinman & Cohn, 1974; Steinman *et al.*, 1974). However, it is more appropriate to talk of a 'family' of DCs because several different types of DCs have been described and they form a heterogeneous population of single cells found in different sites of the body. However, all DCs perform a similar important role, the processing and presentation of antigens leading to the activation of naïve T lymphocytes. Steinman showed in 1974 that DCs are at least 100-fold better activators of T cell responses in a mixed leukocyte reaction when compared to other antigen presenting cells such as macrophages and B cells (Steinman & Witmer, 1978). Later on it was demonstrated that DCs are able to induce a T cell antigen specific response both *in vitro* (Nussenzweig *et al.*, 1980) and *in vivo* (Inaba *et al.*, 1990). DCs are also involved in the establishment self-tolerance. Tolerance is mediated by DCs in two ways, by the stimulation of clonal deletion of T cells during antigen presentation in a steady state (Bennett & Clausen, 2007; Hawiger *et al.*, 2001) and by modulating regulatory T cell differentiation (Luo *et al.*, 2007). DCs have an essential function in the immune system participating in the immediate response to pathogens through innate immune activation as well as in the induction of antigen specific immune responses.

DCs are located in lymphoid organs such as the spleen and lymph nodes and in non-lymphoid organs such as the skin or liver. They can be generally classified into two categories: conventional DCs which are blood-borne, and plasmacytoid DC which are differentiated from bone marrow (Ardavin, 2007). Conventional DCs are present as immature or mature APCs, resident in lymphoid organs such as the thymus, spleen, lymph nodes, lymphoid tissues and the intestinal and respiratory tracts. They can be separated into subpopulations based on their differential expression of the surface markers CD8, CD11b, CD4. Conventional immature DCs that are located within non-lymphoid organs are able to migrate to the lymph nodes upon contact with external antigens, where they interact with antigen specific T cells and other cells of the immune

system like NK cells, B cells and other DCs. This DC population include Langerhans cells and are also called tissue-resident DCs. Similar DCs are also present in the parenchyma of the liver and kidney. These cells are strategically located in intimate connection with epithelial surfaces that are exposed to external pathogens. Tissue-DCs form a subpopulation expressing specialized endocytic and phagocytic receptors that permit proficient uptake and procession of pathogen-specific antigens. However, in the last few years it has been established that the involvement of tissue-derived DC in priming antigen-specific T cells after microbial infection of the skin or mucosa is limited. The arrival of dermal DC in the draining lymph nodes coincides with antigen presentation by lymph node-resident DC, suggesting that dermal DC transport the antigen cargo to resident DC in the lymph node, which in turn activates specific T cell responses (Allan *et al.*, 2006). Plasmacytoid DCs are specialized in producing large quantities of type I IFN, particularly during viral infection. They are non-conventional DCs with low expression levels of CD11c, CD11b and MHC class II markers but they are $B220^{+}$. In the steady state, they resemble plasma cells and upon encounter with antigen they assume a DC characteristic morphology with weak property of antigen presentation (Liu, 2005).

Exploiting the functions of different DC subsets *in vivo* is challenging since the depletion of only one type of DC at a time can present technical problems due to subtle differences in marker expression and overlapping functions between subsets. However, a recent report investigating DC-ablated mice highlighted the crucial role of DC in generating anti-microbial T cell immunity (Bennett & Clausen, 2007).

1.6.2 Antigen interaction

Pathogens are recognized by specific pattern recognition receptors (PRRs) which include transmembrane toll-like receptors (TLRs), cytosolic nucleotide oligomerisation domain (NOD)-like receptors (NLRs), RIG-1-like receptors (RLRs) and C-type lectin receptors (CLRs). Innate immune cells such as macrophages and DCs sense the pathogens through TLRs, phagocytose them and evoke immune responses. The recognition of pathogen-associated molecular patterns (PAMPs) is crucial to the host in order to avoid the activation of innate immunity against self (Takeuchi & Akira, 2007). To date ~13 different TLRs have been described in mice and ~11 in humans. Each TLR

recognizes distinct PAMPs derived from bacteria, viruses, protozoa and fungi (Figure 1.5). TLR1, 2, 4 and 6, which are expressed on the cellular membrane, recognize a broad range of microbe-derived lipid structures; TLR5 and 11 are specialized in protein recognition; TLR3, 7, 8 and 9 are localized intra-cellularly in the endoplasmic reticulum (ER) or endosomal membrane where they detect nucleic acids from viruses or bacteria (Akira *et al.*, 2006).

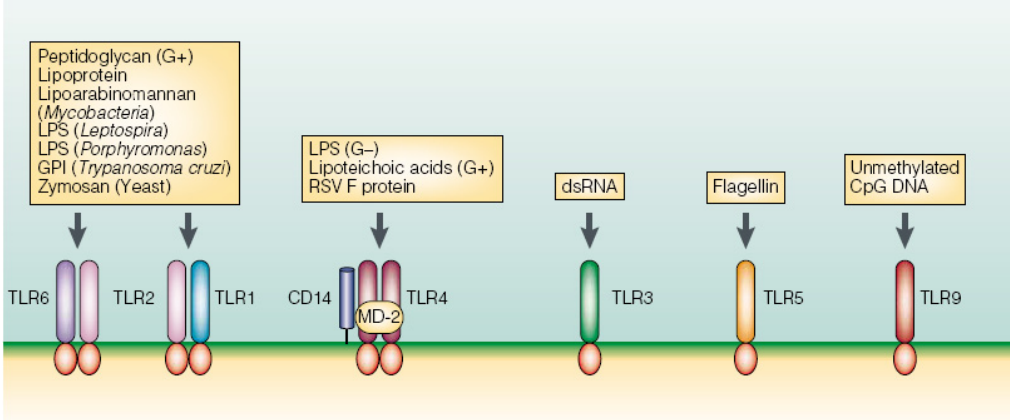

Figure 1.5 TLRs and their respective binding products

This image summarizes some of the TLRs found in mice and the respective PAMPs that they recognize. The best characterised of the TLRs is perhaps TLR4 which recognizes and binds to LPS with the help of two accessory proteins CD14 and MD-2; TLR2 binds a broad range of structurally and functionally unrelated ligands as hetero-dimers with other TLRs like TLR1 and 6; TLR3 recognizes unfamiliar double-stranded DNA; TLR5 specifically binds to bacterial flagellin and TLR9 reacts with unmethylated CpG motifs. G+, gram-positive, G-, gram-negative, GPI, glycophosphoinositol, RSV, respiratory syncytial virus (Medzhitov, 2001).

TLRs are normally expressed as homo- or heterodimers of transmembrane proteins. The extracellular domain is composed of leucine rich repeats that participate in the recognition of PAMPs. The intracellular domains initiate the signaling response through the Toll/interleukin-1 receptor (TIR) domain that interacts with TIR-domains containing adaptor molecules. Four principally different adaptor proteins have been described so far, myeloid differentiation primary response gene 88 (MyD88) is utilized by all the TLRs except TLR3 which transduces a signal through TRIF. Some TLRs use the adaptors Mal or TRAM to bind to MyD88 or TRIF. The majority of TLRs signal through the MyD88 adaptor protein leading to a pathway that involves the molecule

TIR-domain containing adaptor protein (TRIF or TICAM1). These complexes interact with IRAK members and TRAF6 leading to the activation of MAP-kinases and nuclear factor-κB (NF-κB)-dependent cytokine secretion. These events can result in the production of pro-inflammatory cytokines such as TNF, IL-1 and IL-6. In the same pathway inhibitors of NF-κB kinase (IKK) complex can be induced (Trinchieri & Sher, 2007). Figure 1.6 reports a few examples of the pathways triggered by TLRs.

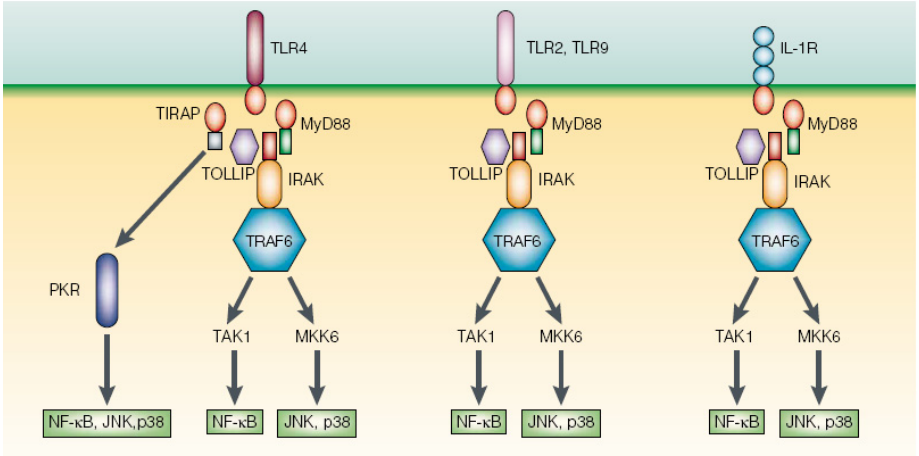

Figure 1.6 TLR induced signaling pathways

This image illustrates examples of the pathways triggered by TLRs. The TLR and the IL-1 receptor family members share signaling machinery such as the adaptor MyD88; Toll-like interacting protein (TOLLIP) has a down-regulating effect on TLR pathway activation; the protein kinase IRAK (IL-1R associated kinase) and TRAF6 (TNF receptor associated factor 6) (Medzhitov, 2001).

Among the non-TLR PAMP binding receptors the NLR family has ~30 members in humans and is characterized by the presence of a NATCH domain and leucine-rich repeats. Only two members of this family have been characterized, NOD1 and NOD2, which bind to two different substructures of bacterial peptidoglycan leading to the activation of NF-κB. Ligand recognition of NLRs, IpaF and NALP3 promotes the assembly of the inflammasome and subsequent caspase-1 mediated activation of the active form of IL-1β and IL-18 and IL-33 (Franchi *et al.*, 2008). Other PRRs include RLRs (RIG-1-like receptors), which are located in the cytoplasm, and some members of this family like TIG-1 and MDA5 mediate the detection of viral RNA resulting in the production of type I IFN (Bowie & Fitzgerald, 2007). Finally the CLR family represents surface expressed PRRs that recognize glycans.

1.6.3 Antigen processing and presentation by DCs

DCs are crucially required for the priming of antigen specific T cell responses, whereas the major histo-compatibility complex (MHC) class I and II presentation pathways are responsible for the priming of CD8+ and CD4+ T cells respectively. These pathways represent major check points for promoting the induction of adaptive immunity directed against intracellular and extracellular pathogens. Pathogens, including some bacteria, can induce host cell apoptosis and in such cases their antigens can engage with DCs upon the phagocytosis of apoptotic bodies (Albert *et al.*, 1998a; Albert *et al.*, 1998b).

MHC class I is responsible for the processing and presentation on the cell surface of antigens derived from endogenous proteins found in the cytosol. MHC class I presents peptide antigens of ten amino acids to CD8+ T lymphocytes. This mechanism seems to be always active in the cell and almost all cells express an MHC class I complex. Cytosolic proteins are ubiquitinylated to mark them for proteolytic degradation in the proteasome complex or by cytosolic enzymes. Once the immune response is activated, three proteosomal subunits are replaced by the 'immunological' homologues LMP2 and LMP7 and MECL1, forming the immunoproteosome complex. The trafficked peptides are translocated into the ER lumen by transporter proteins involved in the antigen processing (TAP). In the ER lumen the peptides undergo N-terminal trimming by aminopeptidases such as endoplasmic reticulum aminopeptidases (ERAP)-1 and ERAP-2. Eventually, the optimal length peptide is associated with the presentation groove of MHC class I. This involves the intervention of specific chaperone proteins that coordinate the assembly of a trimeric complex containing MHC class I heavy chain, β2-microglubulin and antigenic peptide. Consequently the MHC class I-peptide complex moves along the secretory pathway to the cell surface (Jensen, 2007).

The MHC class II complex presents exogenous peptide to CD4+ T cells. The association of the MHC class II complex with the peptide antigen is linked to the endocytic route and endosome/lysosome proteolysis. The formation of the MHC class II complex take place in the ER lumen where the α and β MHC class II chains assemble with a transmembrane protein called the invariant chain (Ii) that directs the formation of large complexes of three MHC dimers and three Ii proteins. Ii is also responsible for the trafficking of the complexes from the ER to the endocytic pathway through the Trans-

Golgi Network (TGN). Newly synthesized MHC class II-Ii complexes accumulate in the late endosome and lysosome. Once fused with the endosome/lysosome the Ii chain is degraded and peptide exchange is catalyzed by the chaperone H2-DM whose activity is triggered by acidic pH. Subsequently, Ii is completely degraded by the cysteine protease Cathepsin S, permitting the MHC class II complexes to reach the cell surface through tubulo-vesicular structures present in mature DCs.

Figure 1.7 reports a simple schematic representation of the events leading to antigen presentation.

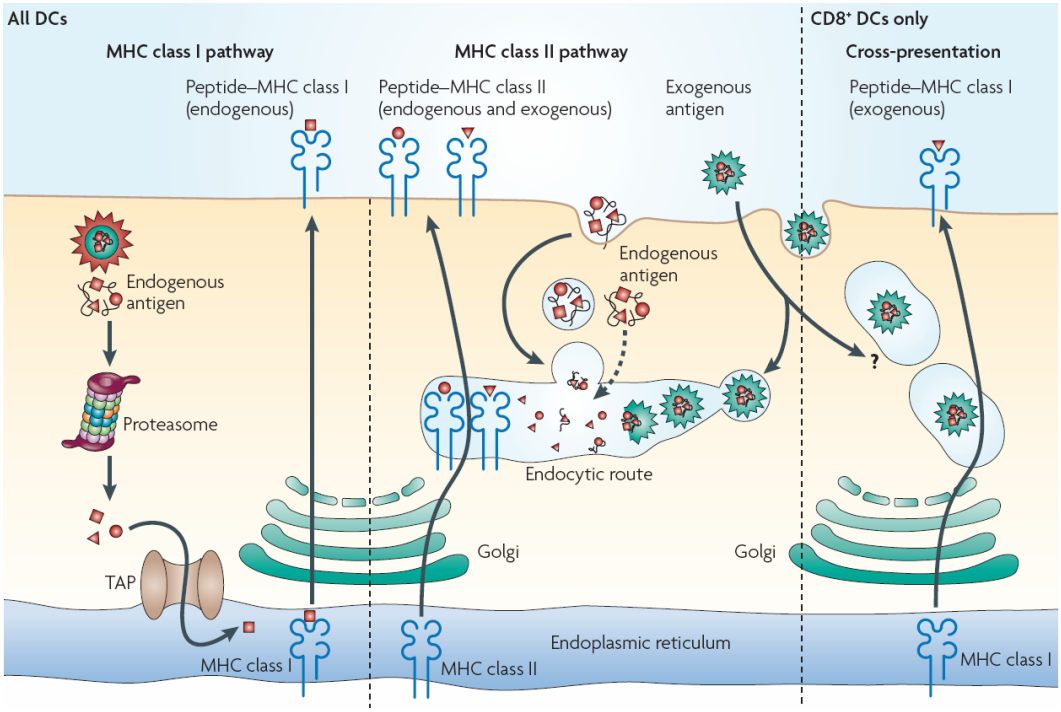

Figure 1.7 Antigen processing and presentation trafficking inside DCs
(Villadangos & Schnorrer, 2007).

MHC class I complexes can present exogenous proteins to CD8+ T cells in a mechanism called cross-presentation. It is not totally clear how this pathway works but it may involve both TAP-dependent and independent mechanisms. The TAP-independent mechanism appears to involve peptide exchange on recycling MHC class I molecules in endosomes or on the cell surface involving Cathepsin S (Shen *et al.*, 2004). The TAP-dependent pathway involves the transport of antigens from endosomes to the cytoplasm and cross-presentation is suggested to result from an adaptation of

retrotranslocation mechanisms dependent on Sec61, involving the degradation of misfolded proteins in the ER (Ackerman *et al.*, 2006). However this antigen presentation pathway has not been fully elucidated yet. Whatever, the exposure of antigen peptide on the MHC class I and II is not sufficient to activate T cells. The co-expression of co-factors is required and T cell activation depends upon their presence. Such co-factors include B7-1 and B7-2 on the DC and on the T cell the CD28, CD4 or CD8 molecule together with the T cell receptor (TCR) (Green, 2000; Greenwald *et al.*, 2005; Yang & Wilson, 1996).

1.6.4 Interaction of DCs with *Salmonella*

Bacteria infecting the gut epithelium preferentially pass through M-cells. However, the ability of DCs to elongate between the epithelial cells and sample the lumen content can facilitate the uptake of bacteria. In this way DCs are suspected to facilitate the penetration of bacteria and if the cell is not able to kill the pathogens they inadvertently become transporters through the epithelial barrier (Rescigno *et al.*, 2001). DCs also contribute to 'immune correspondence' with the lymph node and more DCs are recruited as the infection progresses (Sundquist & Wick, 2005). Generally DCs have some capacity to kill micro-organisms and they express on their cell surface specific antigens such as MHC class I and class II that participate in the initiation of adaptive immune responses. Potential interaction between *Shigella* and DCs has not been well investigated whereas *Salmonella*-DC interaction has been used as a model to study the interaction between invasive pathogens and antigen presenting cells. *Salmonella* demonstrate unusual intracellular behaviour inside bone marrow derived DCs (BMDCs). *Salmonella* reside inside DCs in a static, non replicating form even though SPI-2 expression is detected (Jantsch *et al.*, 2003). In addition, intracellular *Salmonella* reduce the ability of BMDCs to present model proteins and stimulate the proliferation of antigen specific T cells (Cheminay *et al.*, 2005). There is also reduced expression of MHC class II on the surface of DCs during *Salmonella* infection (Mitchell *et al.*, 2004). These studies have reported interference by the SPI-2 TIIISS on the ability of DCs to present antigen on their surface and advanced the hypothesis that SifA may be interfering with cell vacuole trafficking.

1.7 Microarray technology

1.7.1 Overview

The microarray is a key component of an *in vitro* platform technology that facilitates the simultaneous expression analysis of tens of thousands of genes. Using microarrays, in a few days it is possible to know which genes are differentially expressed under different growth conditions or within tissues, for example comparing normal to cancerous cells. Microarray analysis requires some knowledge of the DNA sequence of the target organism and is dependent upon sequence-specific probes representative of each gene. DNA templates are reproduced and bound to a surface (this can be glass, beads or other substrates) generating a genome array. Many copies of the same probes are spotted in an ordinate position and in the size order of less than a micron. This makes it possible to have tens of thousands of spots on the same small matrix. Then, RNA extracted from a target biological sample is labelled using various biochemical tags that can be detected by fluorescence (cyanine dyes), colorimetric detection (biotin/enzyme) or radioactivity. The labelled-RNA is incubated with the microarray slide under conditions where individual RNA molecules will bind to complementary single strands spotted on the array. Any un-bound labelled-RNA is then washed away and bound RNA is revealed by a scanning procedure that senses and records the signal emitted by the labelled-RNA. The intensity of the hybridized probes is measured to estimate the abundance of each target transcript.

The concept of DNA arrays was already in use in the 1970s in the form of dot and slot blots, where radioactive labelling was used to measure the relative expression of a few genes in a target sample (Kafatos *et al.*, 1979). However, the development of microarrays as we know them today has been strongly improved by the automation of PCR, sequencing technologies, progress in robotics, imaging systems and in computational biology (Ewis *et al.*, 2005). The development of new technologies promises to introduce more powerful expression analysis tools. For example, high throughput sequencing machines, that has a reported accuracy of over 99.9%, will improve the precision of interrogation of transcript levels (Margulies *et al.*, 2005). Moreover the development of nanoarrays promises to overcome the microarray

technical problems of ‘point-of-care’ and ‘field applications’. Nanoarrays are already proposed to be suitable for label-free nucleic acid analysis, for protein detection and for enzymatic-based assays with further potential applications in drug discovery, medical diagnosis, genetic testing, environmental monitoring and food safety inspection (Chen & Li, 2007).

1.7.2 Microarrays and gene expression

DNA microarray technology has been mainly applied to quantify mRNA expression or determine single nucleotide polymorphisms (SNPs). In biomedical applications, microarrays are providing enormous amounts of information on genotype-phenotype relationships, gene-environment interactions, helping to understand diseases, aging, mental illness, diet, drug and hormone effects. It is possible to infer molecular phenotyping of diseases through gene expression profiling revealing genes up- or down-regulated in a given physiological state. This can provide diagnostic, prognostic and mechanistic insights that improve our understanding of human diseases. Microarray analysis offers the opportunity to investigate complexes diseases where multiple factors are affecting the illness outcome, like in cancer. Microarrays provided invaluable tools to classify and understand the mechanisms of induction and progression of different forms of cancer. The use of microarray in cancer studies highlighted that different clinical cancer phenotypes can correlate with different genetic characteristic and the analysis of expression profiles can give an indication of the disease grade, clinical course and response to treatment or the propensity to metastasis formation. All these factors come together to provide useful information for cancer prognosis.

Microarray technology is finding applications in drug discovery, investigating the impact of a medication on general genome expression patterns and defining potential side effects. An important goal of modern medicine is to achieve a personalized drug dose for each person depending on their responsiveness to the drug. More importantly, microarrays can give insight into the molecular profile of different cell types under normal functioning conditions. For example, such analysis can help in the understanding of developmental events, or highlight genes specifically expressed in a tissue. This perhaps will help to understand different physiological mechanisms and identify specific genes expressed in different loci of the body. For example, Ewins

reported in 2005 a study of expression patterns in 10 different human tissues (Ewis *et al.*, 2005). Microarrays will also find future application in proteomic research, looking at key interactions between proteins and elucidating complicated pathways involved in the different biological processes. In summary, microarray are being applied in many different fields. However, a future bottleneck may be in data analysis and interpretation. Moreover microarray technology is still facing strong criticism from some quarters. A general lack of standardization of methods and procedures makes it difficult to confirm and correlate observations by different groups. For an exhaustive report of the subject please refer to Ewis 2005 (Ewis *et al.*, 2005).

1.7.3 Other uses of microarrays

Microarray technology is currently utilized in many different areas of research from applied to basic study and from medical diagnostics to prognostics. Microarrays have also been applied in genome wide association studies to identify SNP variation associated with complex disorders. The International Haplotype Mapping Project and Perlegen Science reported more than 2 million SNP markers (Syvanen, 2005). However, it has been estimated that the human genome includes ~10 million hot spots of variation and to map genes associated with diseases can require the genotyping of hundreds or thousands of SNPs. Many commercial microarrays are available for this type of study and they differ in terms of the number of SNPs analyzed and the number of samples that can be loaded per chip (Syvanen, 2005). Also microarray based techniques are available for genome wide mapping of protein-DNA interactions and epigenetic markers. In order to identify protein-DNA binding sites a frequently applied technique is chromatin immuno-precipitation (ChIP) combined with microarray analysis (ChIP-chip). An example of microarrays applied to the investigation of methylation sites is the use of the McrBC enzyme that cuts only methylated DNA. Several variations to this approach are available (van Steensel, 2005). Such research will help to discover mechanisms of gene expression regulation and will provide the knowledge to intervene in deregulated systems. Also, microarrays are finding applications in research on small-RNAs and RNAi cell microarrays are being employed in ‘loss-of-function’ analyses (Wheeler *et al.*, 2005)

1.7.4 Analysis tools

Microarrays representing whole genomes can produce massive amounts of data and the nature of the analysis is crucial to the questions the researcher is aiming to address. The same data can be analyzed with different programs and different aspects of cell homeostasis can be investigated. Usually microarray expression studies are carried out to compare two or more samples with one used as control sample representing the 'normal status' (a reference sample). The control or reference sample is normally compared to a 'treated' or 'different' sample in order to comparatively investigate any differences in the gene expression. Data from replicate samples is combined and the output depends on gene annotation of the specific genome and statistical tools used in the analysis.

As a prelude to any analysis the recorded signals from each single array must be normalized before any comparisons are made. This practice is called data pre-processing and includes imputing missing values, normalization and necessary actions are specified depending on the analysis tools used. Normalization is performed in order to minimise systematic technical errors that are a risk of variability in the microarray employed. All 'microarray chips' have a background signal to which all the single spot signals need to be normalized. Further, chips need to be normalized to each other before they are compared. The normalization of the signal intensities can be done using different methods. In this study GCRMA was used for array normalization, which stands for Robust Multi-array Average with GC-content background correction taking into account the background of the chip and the intensity recorded for the perfect match (PM) and the mismatch (MM) probes based on their sequence information (Wu *et al.*, 2004). Once the normalization is performed the arrays can be compared using different statistical packages.

Pair-comparison analysis highlights genes that are repressed or activated in one sample compared to the other giving fold changes in expression. The differential list can be further filtered utilising statistical analysis of the samples based on significant p-values. For this reason it is important to run at least three biological replicates otherwise a robust statistical analysis cannot be performed. The data can be filtered with different statistical formulae, for example Benjamini-Hochberg False Discovery Rate or

ANOVA. The data analysis can be performed using computer programming language like R or using graphical user interface packages for non-bioinformaticians like GeneSpring or GEPAS (Montaner *et al.*, 2006). Gene lists can be further analyzed by clustering or pathway analysis. Clustering can be a powerful approach to identify gene expression 'signatures' associated with a certain disease or cancer. This method analyses individual genes taking into consideration only those genes differentially expressed above or below a certain threshold. Pathway analysis aims to identify pathways affected during treatment. All these steps will produce different data sets and this must be taken in consideration during their interpretation. Methodologies should be appropriate for the samples under consideration and the biological question addressed.

1.8 Aims of the thesis

The aims of this study were to investigate new approaches to studying host-pathogen interactions using *in vitro* models. This study focused on the host response to two Enteropathogenic bacteria, *Shigella flexneri* and *Salmonella enterica* serovar Typhimurium. In the first part of the thesis a new *in vitro* cell model was employed, mouse embryonic stem cells. Since this *in vitro* model is new to the investigation of pathogenic bacteria firstly it was investigated ES cell's interactions with the pathogens. This enabled to identify if the bacteria behave as previously described in studies using *in vitro* differentiated cells. Subsequently, transcriptomic profiling of murine ES cells during *S.* Typhimurium infection at 2h and 4h incubation was performed. Whole genome arrays were employed in order to investigate the cells'reaction to the bacterial invasion with the view to highlighting patterns of interaction. Additionally, murine ES cells were driven to differentiate into antigen presenting or dendritic cells. The differentiation protocol was optimized for the murine ES cell line used in the first part of the study and the derived esDC were characterized for surface markers and in functional assays. Lastly, the esDC transcriptome active during *S.* Typhimurium infection was investigated revealing novel insights into host-pathogen interactions.