Functional studies of the *PARK2/PACRG* leprosy susceptibility factor

Aimée-Angélique Bouka

Department of Human Genetics

McGill University, Montreal

April 2013

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

Abstract

Background

Leprosy is a chronic human infectious disease that is caused by *Mycobacterium leprae*, a slow growing intracellular parasite mostly of macrophages and Schwann cells. In 2010, there were an estimated 228,474 new cases worldwide. Research conducted for decades has strongly suggested that genetic factors participate in host susceptibility to leprosy. Employing a positional cloning approach, the Schurr group had identified genetic variants in the shared *PARK2/PACRG* promoter region as major leprosy susceptibility factors. Specifically, it was possible to identify two *PARK2* promoter allelic combinations that were strongly associated with leprosy susceptibility. The reference allelic combination that was not associated with leprosy was considered a "resistant" combination while the one which was associated with leprosy (OR = 5.28; CI 95% = 2.06–13.55) was considered a leprosy "susceptible" allelic combination.

Results

Using two strains of transgenic mice carrying a human *PARK2* promoter overlapping the leprosy risk factors fused to a firefly gene reporter, we studied the activity of the human *PARK2* promoter and the leprosy susceptibility alleles in mice. We found a tissue distribution of the reporter construct in mice that was consistent with the expression of *PARK2* in human organs. However, we also noted that the transgenes had poor correlation with endogenous mouse *Park2* expression. Unexpectedly, pPARK2/FLuc transgenes were down-regulated in the spleen of mice following BCG and *S. Typhimurium* infection. Likewise,

endogenous mouse *Park2* expression was repressed in the spleen by BCG and *S. Typhimurium* infection as well as by LPS exposure. We did not detect consistent differential allelic expression of the two pPARK2/FLuc transgenes.

Conclusion

Parkin expression was modulated in the spleen by immune stimulation. This finding added a new element supporting the hypothesis of Parkin being a host defense protein. The transgenic mouse model did not provide data that supported a direct role of the human leprosy susceptibility factors in *PARK2* expression levels.

Résumé

Introduction

La lèpre est une maladie chronique infectieuse affectant les êtres humains. Elle est causée par *Mycobacterium leprae*, un parasite intracellulaire à croissance lente ayant un tropisme pour macrophages et les cellules de Schwann. En 2010, il y eu environ 228.474 nouveaux cas de lèpre diagnostiqués dans le monde entier. Des recherches menées depuis des décennies ont fortement suggéré que des facteurs génétiques de l'hôte contribuent fortement à la susceptibilité à la lèpre. Utilisant une approche de clonage positionnel, le groupe Schurr a identifié des polymorphismes génétiques dans la région promotrice partagée par les gènes *PARK2* et *PACRG* comme facteurs prédisposant à la lèpre. Plus précisément, il a été possible d'identifier deux combinaisons alléliques principales localisées dans le promoteur du gène *PARK2* qui étaient statistiquement associées à la lèpre. La combinaison allélique sans association à la lèpre était considérée comme étant la combinaison conférant une «résistance» contre la lèpre tandis que celle qui a été associée à la maladie (OR = 5,28, IC 95% = 2,06 à 13,55) a été considéré comme étant la combinaison conférant "susceptibilité" à la maladie.

Résultats

En utilisant deux souches de souris transgéniques contenant un promoteur humain du gène *PARK2* ainsi que les facteurs de risque pour la lèpre fusionné avec un gène rapporteur codant pour la luciférase, nous avons étudié l'impact des deux combinaisons alléliques liées à la lèpre sur l'activité du promoteur humain*PARK2*. Il a été déterminé que l'expression du rapporteur chez la souris en

terme de distribution tissulaire est comparable à l'expression de *PARK2* chez l'homme, mais que l'expression des transgènes a une faible corrélation avec l'expression de *Park2* endogène chez la souris. En outre, les transgènes pPARK2/FLuc sont régulés à la baisse dans la rate par lors d'une infection avec BCG et *S. Typhimurium*. De plus, l'expression endogène de *Park2* chez la souris est réprimée dans la rate par l'infection au BCG et à la *S. Typhimurium* ainsi que l'exposition au LPS. Enfin, nous avons pu conclure qu'il n'y a pas d'expression différentielle stable entre les deux allèles du transgène pPARK2/FLuc.

Conclusion

L'expression de Parkin est modulée dans la rate par la stimulation immunitaire. Cette observation ajoute un nouvel élément qui soutient que *PARK2* appartient au mécanisme de défense de l'hôte. Nos lignées de souris transgéniques n'ont pas apporté de résultats prouvant l'impact des facteurs de susceptibilité de la lèpre sur l'expression différentielle du promoteur *PARK2*.

Acknowledgments

First of all, I would like to thank my supervisor Dr Erwin Schurr for giving me this unique opportunity to be introduced to the scientific world in an excellent environmental setting and for assuring my financial support. I want to thank him for his insight and advices. He led me throughout my studies and helped me develop critical thinking and originality which are both necessary in science.

I want to offer a special and sincere thank you to Dr Tania Di Pietrantonio for teaching me all the basics of various procedures as well as analysis and for providing crucial technical assistance with mouse work and quantitative PCR. Her help and support were priceless and I will be forever grateful. Thank you for motivating me to be better and for nurturing such a nice friendship with me. Nothing would have been the same without you.

I also want to thank the other members of my lab for their kind support and their presence. Dr Marianna Orlova, thank you for your dedication and your much needed assistance at a critical time, but mostly thank you for understanding me and for feeding my insatiable stomach. Jason Yu, thank you for your cheerful company. Dr Vinicius Fava, Dr Jeremy Manry and Dr Aurelie Cobat, thanks for making the lunchtime the best time of the day, for making the lab such a party and also for making human genetics easier to understand. Maylis De Suremain and Laurel Holmes, thank you for being for bringing so much warmth to the lab.

Finally, I want to thank members of my family that gave me with encouragement and support throughout my Master's degree. I want to offer a special thanks to my parents Papounet and Mamounette for always believing in my abilities and my sister Yolande for being my number one fan. Merci du fond du coeur. At last, I want to thank my dear husband for his love and understanding, but mostly for thinking that I am a genius. Thanks peanut for giving me the motivation to finish in due time. A little pressure does not hurt ;)

Table of Contents

List of tables and figures

Table

Figures

List of Abbreviations

AIDS	Acquired immune deficiency syndrome
APC	Antigen presenting cell
BB	Borderline borderline leprosy
BCA	Bicinchoninic acid
BCG	Bacille Calmette-Guérin
BL	Borderline lepromatous leprosy
BT	Borderline tuberculoid leprosy
CCDC122	Coiled-coil domain containing 122
CFU	Colony forming unit
CO_2	Carbon dioxide
DNA	Deoxyribonuclase acid
E1	Ubiquitin-activating enzyme
E2	Ubiquitin-conjugating enzyme
E3	Ubiquitin-protein ligase
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GWAS	Genome-wide association study
HLA	Human leukocyte antigen
Hprt	hypoxanthine phosphoribosyltransferase 1
IBR	In between RINGs
IL-12	Interleukin 12
IL-2	Interleukin 2
IL-4	Interleukin 4
INFγ	Interferon gamma
IP	Intraperitoneal
IV	Intravenous
LACC1	Laccase (multicopperoxidoreductase) domain containing 1
LD	Linkage disequilibrium
LL	Lepromatous leprosy
LPS	Lipopolysaccaride
LRRK2	Leucine-rich repeat kinase 2
MDT	Multidrug therapy
MHC I	Major histocompatibility complex I
MHC II	Major histocompatibility complex II
NO	Nitric oxide
NOD2	Nucleotide-binding oligomerization domain-containing protein 2
PAEL-R	Probable G-protein coupled receptor 37

PACRG	Parkin Co-regulated
Park2	Parkin (mouse)
PARK2	Parkin (human)
PBS	phosphate buffer solution
PGL-1	phenolic glycolipid-1
PINK-1	PTEN-induced putative kinase 1
pPARK2/FLuc	PARK2 promoter and firefly luciferase
RBR	two RING finger domains separated by an IBR domain
RING0	Really interesting new gene 0
RING1	Really interesting new gene 1
RING2	Really interesting new gene 2
RIPK2	Receptor-interacting serine/threonine-protein kinase 2
RLU	Relative light unit
RNA	Ribonuclease acid
ROI	Reactive oxygen intermediate
RT-PCR	Real-time polymerase chain reaction
Slc11a1	Natural resistance-associated macrophage protein 1
SNP	Single nucleotide polymorphism
TH1	T helper 1
TH2	T helper 2
TNF	Tumor necrosis factor
TNFSF15	tumor necrosis factor (ligand) superfamily, member 15
TT	Tuberculoid leprosy
UV	Ultraviolet
WHO	World Health Organization

Rationale, hypothesis, objectives

Rationale

A large-scale study of leprosy susceptibility in Vietnamese and Brazilian patients identified polymorphisms in the shared promoter region of the *PARK2* and *PACRG* genes as global leprosy risk factors. Further studies directly implicated Parkin in leprosy susceptibility. Hence, we decided to study in a mouse model the effect of a two SNP genotype promoter polymorphism associated with either increased susceptibility or resistance to leprosy in humans. Two strains of transgenic mice were engineered by adding the human *PARK2* promoter region with respectively one leprosy susceptibility allelic combination and a reporter gene using the *Hprt* locus for specific genomic insertion. In this way, the impact of both allelic combinations on the *PARK2* promoter activity in different conditions can be studied across different mouse organs.

Hypothesis

The "resistant" and "susceptible" allelic combinations of the *PARK2* promoter drive different levels of *PARK2* expression either constitutively and/or after infection with bacterial agents.

Objectives

- To confirm the expression of pPARK2/FLuc constructs in target organs using a transgenic mouse model
- To study the impact of infection on pPARK2/FLuc reporter and endogenous *Park2* expression in organs and cell types of interest

- To observe differential expression of *PARK2* promoter variants in a natural state and/or upon infection

CHAPTER 1

Introduction

Leprosy

Historical overview

Although a young Norwegian doctor, Armauer Hansen, was the first to identify and characterize the bacterium responsible for leprosy in the late 19th century[11], the disease had already been haunting and tormenting humankind for centuries. The recent discovery of the earliest confirmed case of leprosy in Jerusalem showed the presence of *Mycobacterium leprae* (*M. leprae*) infection through genomic analysis of the osteological remains of a man who lived in the first century A.D.[12]. Nevertheless, ancient Egyptian texts suggest the existence of the debilitating affliction more than two millennia before Christ[13]. The disease was long thought to be strictly hereditary[14] as well as some sort of curse, stigmatizing afflicted individuals and their family[15]. However, the discovery of the causative bacilli even if contested at first, classified leprosy as an infectious disease[16].

In earlier times, leprosy was treated with Chaulmoogra oil[17]. Not only was the regimen painful, but it also was only partially potent. Only some receiving the treatment were found to be healed and among them, the chance of relapse was high[18]. It was during the Second World War that dapsone, a bacteriostatic antibiotic which inhibits the folic acid synthesis pathway, was introduced as a new cure for leprosy[19]. Unfortunately, less than 15 years later, resistance became an

issue with the administration of this drug[20]. In the 1970s, rifampicin became the curative agent of choice for the treatment of leprosy for its high bactericidal efficiency[21]. Once again, antibiotic resistance became a concern less than a decade after the initial use of rifampicin as a chemotherapy[22]. Clofazimine, which mechanism of action is not fully understood, had been used as a treatment for leprosy since the 1960s[23]. Yet, drug resistance has rarely been reported. Its anti-inflammatory ability is used to treat reversal reaction which occurs in some leprosy affected individuals[24]. Since 1981, the World Health Organization (WHO) has recommended multidrug therapy (MDT) using conjointly dapsone, rifampicin and clofazimine as an effective cure for leprosy[25] and the treatment has been made available free of cost in all endemic countries since 1995[26]. The treatment is followed for a variable amount of time depending on the severity of the disease, typically between 6 and 12 months[27].

Pathophysiology and clinical classification

Leprosy is defined by two main features: the first one being a mycobacterial infection and the second being a peripheral neuropathy[28]. The infection causes a chronic granulomatous inflammation of the skin and peripheral nervous system leading to nerve damage[6]. The symptoms of the disease often include the "famous" anesthetic skin lesions, a loss of sensation and motor abilities and in some extreme cases, shortening of the limbs[29]. Despite the fact that the disease does not cause death, its clinical presentation in its worst cases can be very serious and extremely debilitating.

The severity of the disease depends mainly on the host immune response to the infections. The WHO classified the disease in two major categories. Paucibacillary (PB) leprosy describes the form of the disease where the afflicted individual presents five or fewer skin lesions. Multibacillary (MB) leprosy is where the person has more than five skin lesions[30]. Most experts and leprologists still refer to the Ridley-Jopling five-group system classification to categorize the different clinical forms of the disease[31]. The five-group system segregates leprosy in subtypes accordingly to the host's cellular immunity against *M. leprae* (Fig. 1.1). The severity of the disease will vary depending on the subtype[32]. We find in one extreme of the spectrum, tuberculoid leprosy (TT) and at the other, lepromatous leprosy (LL). Between these two poles are found the intermediate forms of the disease: borderline tuberculoid (BT), borderline borderline (BB) and borderline lepromatous (BL)[31]. The Ridley-Jopling classification includes five main groups, but in reality the immunological scale is graded in more subcategories[32].

The tuberculoid pole is defined by T cell-mediated immunity against the pathogen. The immune response in TT is characterized by a high number of lymphocytes actively participating to the host defense. The granulomatous lesions are epitheloid, meaning that the lesions involve principally macrophages with an "epitheloid" appearance. There are virtually no mycobacteria found in lesion biopsies in individuals of TT leprosy. In this instance, a single well-defined hypopigmented lesion will afflict the patient[28]. Neuropathy however can be severe and occur rapidly in TT. The lepromatous pole is defined by a very high bacterial

load and ineffective cell immunity with significant serum immunoglobulin levels. The granulomas contain large numbers of bacilli inside macrophages. The skin lesions in this case are numerous and highly diffused[28]. The borderline subtypes show clinical and immunological features similar to one pole or the other.[10]

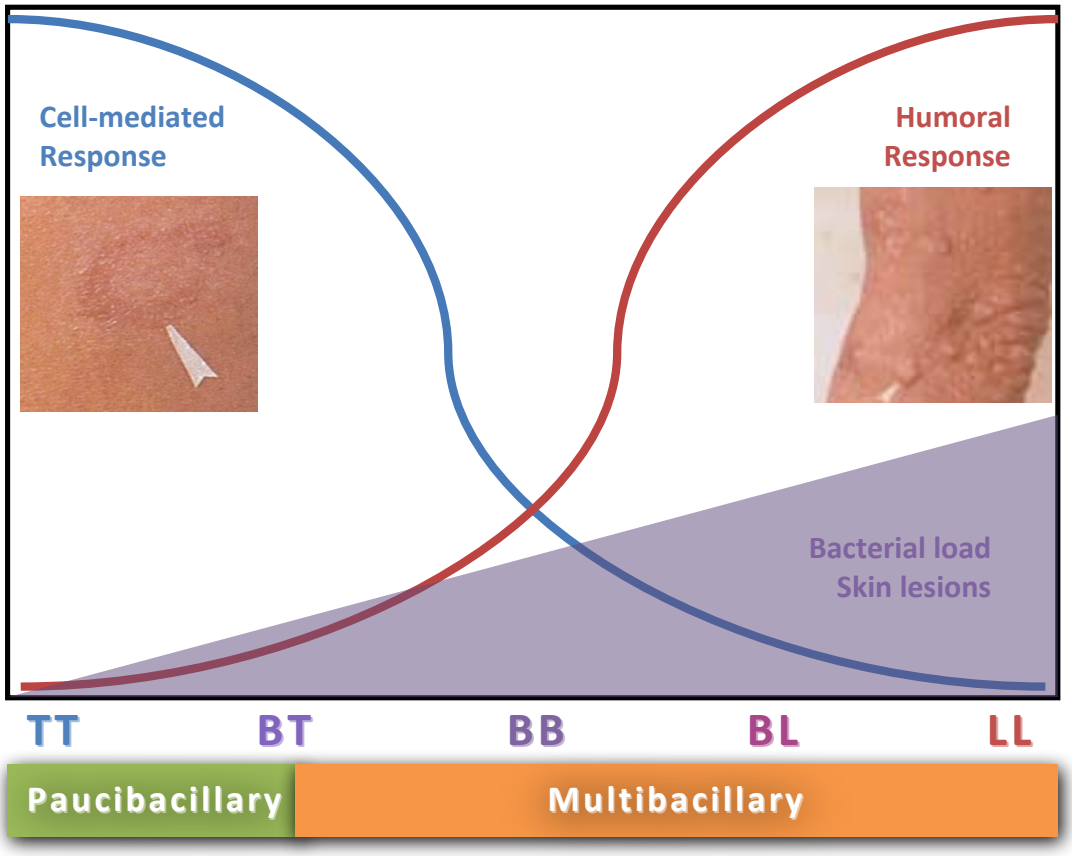

Figure 1.1 Clinical spectrum of leprosy
The clinical spectrum of leprosy elaborated by Ridley and Joplin extends from the least severe pole (TT) characterized by cell-mediated immune response and to the other defined by a humoral response (LL). The immunological spectrum is progressive and a continuum where skin lesions as well as bacterial load increase with the severity of the disease. The paucibacillary form corresponds to TT and some BT cases, the others borderline subtypes and LL are more consistent with multibacillary leprosy[9]. Photographs obtained through http://images.md. laneproxy.stanford.edu/

Pathogenesis

It is generally accepted that leprosy originates from the infection of specific cell types by the pathogen *M. leprae*. The bacteria have a tropism mainly for Schwann cells, which sheath the peripheral nerves, and monocyte-macrophages. Even though the mode of transmission is not clearly understood, it is believed that the disease pathology depends on how these cell types are affected. *M. leprae* is believed to be transmitted via respiratory droplets, and reaches its cellular targets through the lymph and blood vessels. The involvement of the nervous system occurs when Schwann cells phagocytize the pathogen through vacuoles. There, the bacteria are able to survive and replicate and eventually create granuloma since Schwann cells do not have enzymes to destroy the pathogen[10]. The bacteria bind to the Schwann cells through a laminin-binding protein which recognizes phenolic glycolipid 1 (PGL-1) specific to *M. leprae*[33]. Monocytes, once stimulated, become macrophages and may act as antigen-presenting cells (APC) Macrophages engulf and attempt to destroy the mycobacterium, a mechanism that is crucial to the disease outcome.

Macrophages in leprosy function differently depending on the leprosy subtypes. In the case of TT (Fig. 1.2A), macrophages have the ability to fully destroy the pathogen and form competent (APC. Adequate antigenic information will be presented on the cell surface. These APC, when associated with major histocompability complex II (MHC II) proteins, will stimulate an immune response by secreting interleukin 12 (IL-12) which will stimulate CD4+ lymphocytes. The CD4+ lymphocytes will, in return, produce interleukin 2 (IL-2)

and interferon γ (INFγ) and these cytokines will activate new macrophages to develop into epitheloid cells. In the event that MHC I is part of the process, CD8+ lymphocytes will be recruited to destroy the pathogen through apoptosis of infected macrophages[10,34].

In LL (Fig 1.2B), macrophages are unable to fully destroy the *M. leprae*, and are only able to partially lyse the bacteria. In that instance, bacterial phospholipids may still be present leading to the formation of lepra cells. Lepra cells, also called Virchow cells, are large macrophages with foamy cytoplasm containing a large amount of live and degenerated bacilli. Until the Virchow cells start aging, no immune stimulation will take place. The aged lepra cells will be ingested by other macrophages and form new APC with different antigenic information. This cascade of events will, through the expression of interleukin 4 (IL-4), lead through the participation of CD4+ lymphocytes to humoral immunity involving B lymphocytes, plasma cells and anti-*M. leprae* antibodies[10,34].

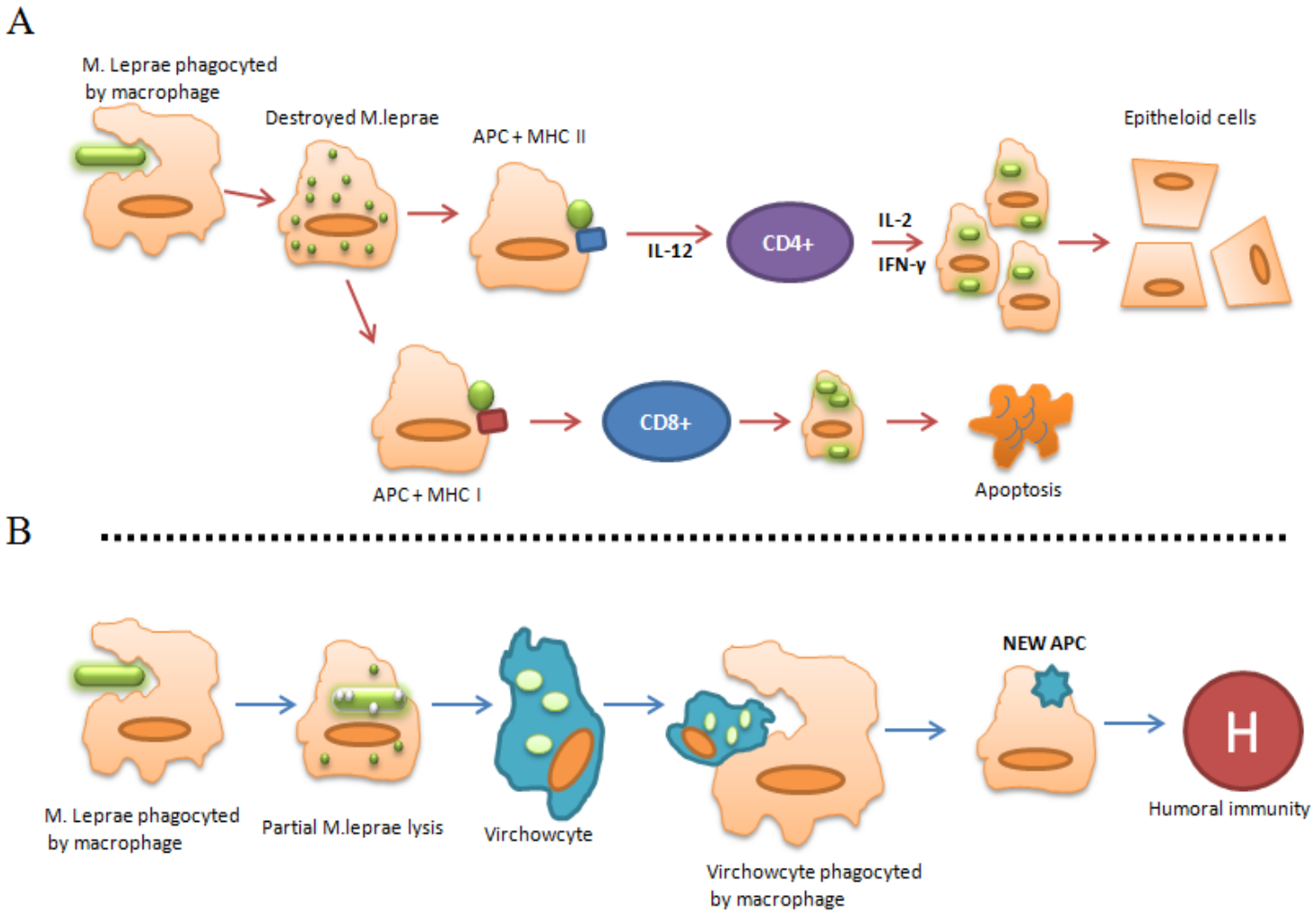
A
M. Leprae phagocyted by macrophage
Destroyed M.leprae
APC + MHC II
IL-12
CD4+
IL-2
IFN-γ
Epitheloid cells
APC + MHC I
CD8+
Apoptosis
B
M. Leprae phagocyted by macrophage
Partial M.leprae lysis
Virchowcyte
Virchowcyte phagocyted by macrophage
NEW APC
H
Humoral immunity

Figure 1.2 Immune reactions in tuberculoid and lepromatous leprosy
In TT (A), Mycobacterium leprae phagocytized by macrophages may be entirely destroyed; appropriate antigenic information may be expressed on the cell surface. In association with MHC class II, these APCs may produce and secrete IL-12 and stimulate CD4+ lymphocytes (Th-1), which subsequently could produce IL-2 and IFN-γ. New macrophages will be activated and transformed into epitheloid cells. When MHC class I is involved, stimulated CD8+ lymphocytes act on other macrophages to kill the organisms by apoptosis. In LL (B), macrophages only achieve partial lysis and the bacterial phospholipids may persist. Virchowcytes may appear and be phagocytized by other macrophages. New APC with modified antigenic information may appear, and stimulate humoral immunity[10].

Epidemiology and transmission

Finding a cure to treat leprosy was a battle that lasted centuries. In the last decades, efficiency was finally attained, but drug resistance remained an important challenge until the mid-80s when multidrug therapy (MDT) was introduced[23]. Though the quest for finding an effective treatment against leprosy was successful, the disease has not been pushed to worldwide elimination. The WHO has defined elimination by a disease prevalence of 1 case per 10 000 inhabitants or less[35]. Even though MDT significantly reduced the prevalence of the disease, the incidence has remained steady (Fig 1.3). More than 200 000 new cases were detected in 2011[36]. In some areas of the world, leprosy is still a major health issue. In fact, it was reported that India, Brazil and Indonesia constituted 83% of the new leprosy cases in 2011[36]. Leprosy is not highly infectious, with may be as little as 5% -10% of individuals exposed to *M. leprae* developing the disease [37]. It is known that a strong genetic component predisposes individuals to contract leprosy.

As mentioned earlier, the exact mechanism of transmission of leprosy is still unknown. However, it is now widely accepted that the main mechanism for transmission of leprosy occurs through nasal droplets to the respiratory system[6]. The possibility that infection occurs through skin contact from affected individuals cannot be excluded[19]. There have also been possible cases of zoonotic transmission of leprosy from armadillos to humans[38]. Some environmental, economic and behavioral risk factors contribute to a higher likelihood of developing leprosy. Household contact with an untreated MB case can raise the risk of contracting leprosy up to 10%[37]. A study conducted in Brazil showed that low education level, bathing regularly in open water, food shortages or the lack of bedding or linen change, were associated with an increased risk of developing clinical leprosy[39].

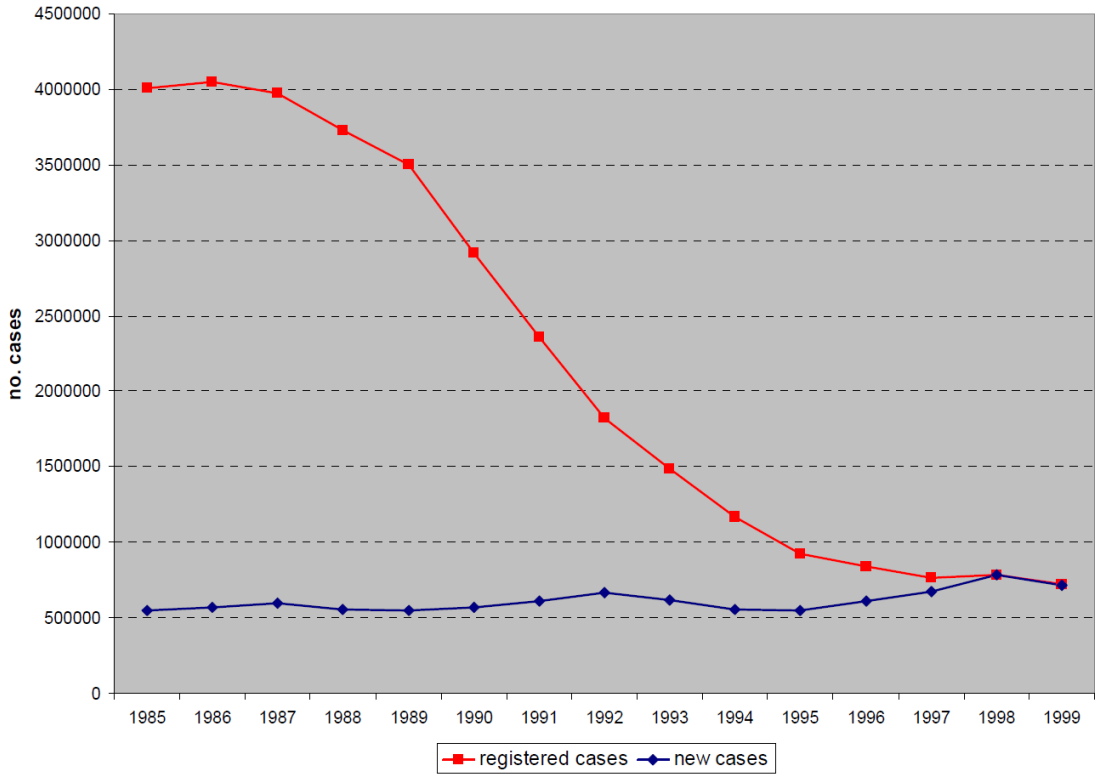

Figure 1.3 Incidence and prevalence between 1985-1999
Even though the prevalence of leprosy has been declining in the number of new cases (incidence) tends to remain stable partially due to the unknown mode of transmission of the disease and the important genetic factor contributing to it[6]. Image from http://www.leprosy-control-studies.net/platform/content/element/2950/2008-10_Input_Study_Design_KIT.pdf

Mycobacterium leprae

Bacteriology

Mycobacterium leprae, the etiologic agent of leprosy, is an acid-fast-staining aerobic bacillus and obligate intracellular pathogen[37]. It was the first human pathogenic bacterium to be identified, preceding the discovery of *Mycobacterium tuberculosis* by a decade[3]. Being part of the Mycobacteriaceae family, *M. leprae* is nonmotile. It is also characterized by the mycolic acid waxy coating layer as well as a lipid-rich cell wall which protects it from desiccation and a great number of antibiotics[40]. The *M. leprae* cell wall is composed of peptidoglycans, arabinogalactan and lipoglycans[29]. The fact that it cannot be cultured *in vitro* is an important challenge in the study of the bacteria[41]. *M. leprae* is a very slow growing mycobacterium which duplicates every 14 days by binary fission[42]. The rod-shaped bacteria tend to arrange like a bundle of cigarettes, never like chains. Moreover, only carbol-fuchsin solidly stained bacilli are thought to be viable, bacilli not as clearly defined by the staining are believed to be dead or dying bacteria. This criteria is clinically important when treating patients[43].

M. leprae is grown in normal and athymic mice footpads for scientific purposes, but except for humans and a few other primates[44], armadillos are the only known natural hosts[45]. Armadillos are thought to be a possible reservoir because of their cooler body temperature[37]. *M. leprae* is unique by its ability to infect nerves by binding of its phenolic glycolipid I (PGL-1) to laminin-2 on the surface of Schwann cells[46].

Reductive genomic evolution

The genome of *Mycobacterium leprae* was fully sequenced in the early 2000s[40]. Soon after, an exhaustive comparative study was conducted using *M. tuberculosis* as a model[4]. It was found that the weaker virulence of *M. leprae* when compared to its cousin *M. tuberculosis* could be explained by a very low coding capacity of the bacillus and its high proportion of pseudogenes[3]. *M. tuberculosis* contains over 4 million base pairs potentially coding for close to 4 000 genes (Fig. 1.4A), whereas *M. leprae,* whose genome is constituted of about 3 million base pairs, only has a coding capacity of approximately 1 600 proteins instead of the estimated 3 000 [4]. These factors significantly affect the metabolic abilities of the bacteria. Only 50% of the *M. leprae* genome sequence is coding compared to 90% for some of its mycobacterium cousins like *M. tuberculosis* (Fig. 1.4B). This low percentage initially created uncertainty about the leprosy pathogen actually belonging to the *Mycobacteriacae* family[40].

The leprosy and the tubercle bacilli, both slow growing mycobacteria, evolved from a common ancestor some thousands of years ago as observed in their phylogenic tree[47]. Despite many similarities between both pathogens, many metabolic systems were left impaired in *M. leprae*. Furthermore, some genes and lost their redundancy in the bacterium's evolutionary process. For example pathways, lipolysis is an important catabolic process in mycobacteria, providing most of their energy through the degradation of the host's lipids. *M. leprae* only has 10% of the lipases compared to *M. tuberculosis*[4]. In summary, the catabolism of the leprosy pathogen is severely limited, with some critical enzymatic

pathways simply absent. The pathogen's genome is a victim of extreme reduction making the organism an obligate parasite, using the host cell metabolism for its own growth and survival.

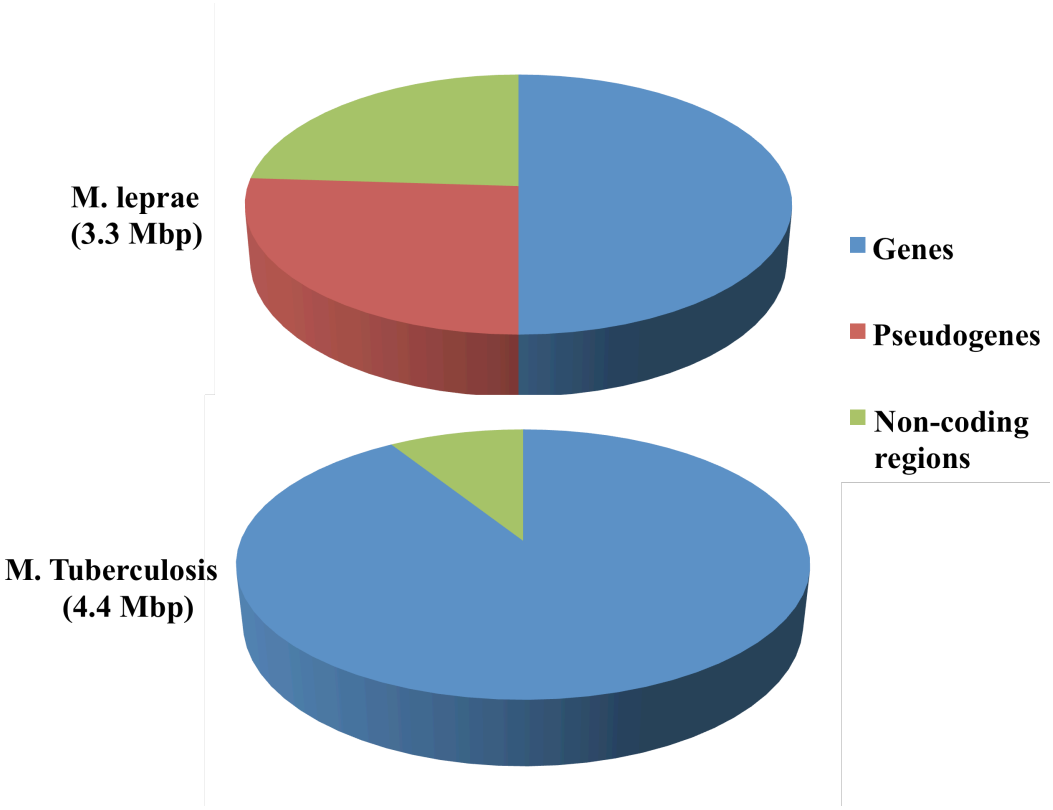

Figure 1.4 *M. leprae* genomic reduction and comparative genomics
M. leprae's genome is close to 3.3 Mbp in size. We can observe an important genomic reduction. Only half of *M. leprae*'s genome encodes for functioning genes. More of a quarter its genome encodes for pseudogenes compared to virtually none for *M. tuberculosis*. *M. Leprae*, through its genomic evolution has become an obligate intracellular pathogen [3,4].

Koch's postulates[48]

Koch's postulates have been used to identify a causative relationship between a parasite and an infectious disease for over a century. However, when applying these proposals to *M. leprae*, it causes, in some cases, an interesting challenge.

1. *The organism or other pathogen must be* ***present invariably in the disease.***

As mentioned previously, LL patients tend to have bacilli-filled lesions, but in the less severe form of the disease, TT, some afflicted individuals do not show any presence of bacilli even if the symptoms concord with a leprosy diagnostic.

2. *The pathogen can be isolated from the diseased host and* ***grown in pure culture.***

Even when disregarding the first postulate, when the pathogen can in fact be found in abundance in the host, *M. leprae* lack too many metabolic pathways to be grown in any type of medium. The use of animal models is required for the purpose of growing *M. leprae*.

3. *The pathogen from the pure culture must* ***cause the disease when inoculated into a healthy, susceptible laboratory animal.***

The leprosy bacillus, when inoculated into a healthy individual, does not necessarily cause the disease. In fact, only an estimated 5-10% of people exposed

to *M. leprae* contract the disease and the 5-10% who do develop disease will be individuals with an immune predisposition. Furthermore, the athymic mouse was genetically engineered to be susceptible to leprosy; therefore the inoculation of *M. leprae* in that animal model would cause the disease.

4. *The pathogen must be* ***reisolated*** *from the new host and* ***shown to be the same*** *as the originally inoculated pathogen.*

If the host is susceptible, it will be possible to reisolate the pathogen from an inoculated host. However, it is possible that a diseased inoculated host may destroy the pathogen through his immune system, but that the individual still be sensitive to some antigens of the bacilli. In that case, the patient might still display symptoms of the disease.

Within the proper hosts, meaning individuals who show the proper susceptibility profile, all but the second postulate may be applicable to the relationship between *M. leprae* and leprosy. Nevertheless, the bacterium is not cultivable in medium and a large number of leprosy cases in human subjects do not correspond to the other three postulates.

Genetics of infectious diseases

Infectious diseases are often seen as a simple relationship between a pathogen and a host. Using the premise that immunity acts in an identical fashion across different individuals, it would seem straightforward to understand the origin or the pathogenesis of a disease caused by an infection. In fact, the host defense system is complex and involves various genes. For these reasons, a great deal of variability can be observed in the reaction of hosts to infection[49]. More than 300 genes in replicated studies have be found to be associated with infectious diseases and commonly these genes encode for chemokines and their ligands, tumor necrosis factors, HLA class II or cytokine receptors although there is a strong bias to study role of such genes in infectious disease susceptibility[50]. Genetic susceptibility to infection is polygenic and multifactorial[49] and is considered a complex trait

Family-based versus population-based analysis

Genetic studies based on related individuals aim to find monogenic and highly penetrant traits[51]. Family-based studies use parents with affected offspring as the basis of the genetic analysis. Both linkage analysis and association testing can be used in family-based studies. Twin studies bring significant strength to family studies. Concordance or discordance of an ailment in monozygotic twins can be very instructive when studying the importance of genetics in infectious diseases[52]. Some advantages of family studies are the discovery of a set of causative genes or pathways for a disease, as well as a greater knowledge or control of genetic and environmental variances[51]. However, they tend to be costly as it is difficult to

recruit related individuals for studies. Moreover, pedigree studies are restrained by the number of generations available. Despite these issues, family-based studies remain a statistically powerful approach to discover genetic variances contributing to infectious diseases.

Population-based studies rely on the case-control approach and generally involve a larger number of individuals. With the elaboration of newer genome analysis technologies, the use of population-based studies has become more and more frequent. The costs of genotyping and sequencing are constantly decreasing, making large-scale population studies more affordable. When a trait is complex, such as the host response to infection, it is advantageous to scan through a population to identify a variety of polymorphisms involved in the host's immunity. However, the situation of genetic background heterogeneity, more specifically admixture, in some populations can confound the results found through population-based studies[50,53].

Linkage studies

Linkage strictly depends on familial genetic transmission and aims to find the physical location of putative disease causing genes in relationship with mapped genetic markers. The markers often are variant polymorphisms in the studied population but always present in the genome. The principle of linkage is based on finding the parental origin of a known marker in linkage with the disease trait. The closer two markers are one to another, the less often recombination occurs; hence the probability of markers being separated by chromosomal recombination is minimal. The term linkage disequilibrium is used to describe a

situation when two markers or region in the genome show a statistically significant pattern of segregation and as consequence display preferential assortment of specific alleles at these loci. A candidate region for a disease gene can be determined when analyzing recombination frequencies against genetic markers across the genome. The causative gene tends not to recombine with markers that are physically close to it.

Linkage analysis has been proven useful to identify disease causative genes where the genes show a strong genetic effect. In the field of genetics of infectious diseases, linkage analysis has led to the identification of some susceptibility genes. However, the results remain scarce[54] and limited to a small number of communicable disease[55]. Linkage analysis through mouse genetics has been successful in research of host susceptibility to tuberculosis[54]. The subject is discussed below.

Genome-wide association studies

Genome-wide association studies (GWAS) rely on a scan of the entire genome in the hope of finding common genetic variants associated with a disease in a population. It is mainly used in a case-control approach. A genomic comparison is made between case and control groups to identify alleles that are significantly found more often in one arm (e.g. cases) then the other (e.g. controls). It aims to create a relationship between variants and a disease without an underlying hypothesis. A variant (single nucleotide polymorphisms are generally used) can be associated with an increased likelihood (susceptibility) or a decreased likelihood (resistance) to develop a disease. Currently, genome-wide

association study is the method of choice for the discovery of genes contributing to host defense to infectious diseases. The HapMap project as well as the improvement of genotyping technology has made genome-wide association studies a weapon of choice to understand the genetic basis of infectious diseases[56]. Despite great promise, GWAS in infectious diseases had only limited success in identifying genetic risk factors of disease.

Human immunodeficiency virus (HIV) was the first infectious disease studied by GWAS and was shown to be a success by the discovery of human leukocyte antigen (HLA) genes *HLA-B* and *HLA-C* responsible for 15% of the variation across individuals for viral load during a specific time in the asymptomatic period of the disease[57]. Thus GWAS was able to identify genetic risk factors not only for infection, but also progression of the disease and response to clinical treatment although both gens had already been identified by candidate gene based studies. Malaria and tuberculosis, known to be among the deadliest infectious diseases were also studied through genome-wide scans but provided disappointing results[56]. However, it has been found that analytic models in genome-wide studies used in the Caucasian population are not as powerful for non-European populations[58]. For example, when a GWAS was performed for malaria, the locus *HBB* known for having an HbS protective allele only presented a weak odds ratio. This only confirms that there is room for improvement for GWAS and computational tools.

An important advantage for GWAS is the ability to discover novel genes and pathways[58]. This approach rarely identifies causative SNPs for a disease, but

instead establishes a correlation between allelic markers or a given genomic region and a phenotypic trait. Despite the possibility of identifying new genetic factors for host susceptibility to a disease, a major downfall for GWAS is that more often than not, the variants identified account for only small consequences of the disease outcome[56].

Candidate gene studies

In the context of genetics of infectious diseases, the candidate gene approach is based on the presumption that certain pathways will influence the host defense against infection, clinical development and severity of a disease. In fact, in candidate gene analysis, the genes to be studied are already predetermined due to knowledge of their biological function[59]. The approach consists of genotyping polymorphisms in relevant genes in both control and affected individuals (in a case – control study) to identify alleles that could explain the phenotypic differences in the host response to a pathogen[60].

This approach has proven to be somewhat successful in the past confirming the involvement of some critical genes in immunogenetics. We can mention here the globins for the resistance to malaria as one example[61]. The human leukocyte antigen can also be mentioned as a gene of importance found to be associated with malaria once again, but also tuberculosis, leprosy, AIDS and hepatitis virus persistence through the candidate gene approach[59].

Moreover, if sufficient numbers of cases and controls are enrolled, candidate gene studies tend to have better statistical power when compared to GWAS[62]. Despite these advantages, a major limitation of this approach is the use

of known genes and pathways as basic hypotheses which makes it impossible to discover novel genes that could partially explain some host immunity differences[62].

Genetics of leprosy

Leprosy has been studied in immunogenetics for a few decades. Hansen discovered much to his surprise that the inoculation of the *M. leprae* bacterium to an individual does not necessarily bring forth the disease. Like many infectious diseases, often the exposure to the pathogen is insufficient to cause disease and an understanding of the host genetics is needed to predict the clinical result[63]. Several genes were discovered to be associated to leprosy whether to partially explain the broad spectrum of the disease or pathogenesis of the ailment.

Stages of leprosy and known genes involved

First, a genetic segregation occurs after exposure to the pathogen. A majority of individual even with constant exposure to the bacteria will never become infected with *M. leprae*. For 90% of the population, contact with the bacilli will never result in a disease. We could assume that some have natural resistance to the mycobacterium. In the event that infection happens, a category of people will undergo asymptomatic clearance of the bacteria[9]. For others, the infection will not clear itself and these individuals will remain asymptomatic carriers. Finally, leprosy per se will manifest itself in the remaining, approximately 10% of exposed people.

We can apply what is called a two-stage model to leprosy (Fig. 1.5). Some genes are involved at different stages of the disease[64]. Among the known genomic

region influencing susceptibility to *M. leprae* infection, the MHC is constantly listed as part of the major ones[58]. The loci *HLA-DR / DQ* were identified as strong susceptibility factors for leprosy. The tumor necrosis factor ligand member 15 (*TNFSF15*) and the coiled-coil domain-containing 122 (*CCDC122*) were among genes found linked to the disease[2].

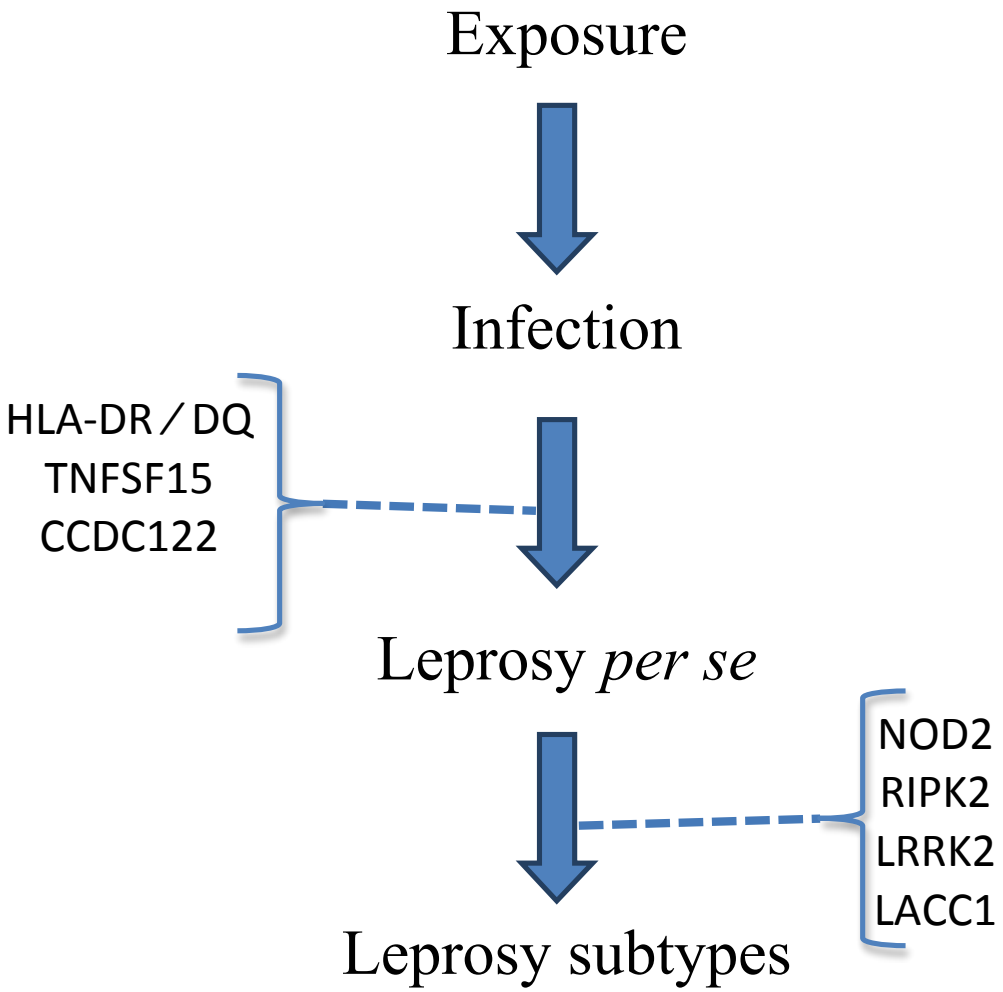

Figure 1.5 Two-stage leprosy model
The genetic contributing factor in leprosy impacts at two steps of the disease. Some genes are known to influence the susceptibility to leprosy itself namely, the *HLA* loci, *TNFSF15, CCD122* and the *PARK2/PACRG* locus. Others genes impact on the clinical subtype and severity of the disease among them *NOD2, RIPK2, LRRK2, LACC1*[2,8].

Polarity of leprosy and host defense

Previously, the two poles and their intermediate clinical manifestations of leprosy were discussed. The importance of the host immune defense was said to be critical to determine the outcome of the disease. Several genes were found to have a stronger association with the more severe MB clinical form of leprosy. Indeed, thenucleotide-binding oligomerization domain-containing 2 (*NOD2*), an important gene in innate immunity as sensor receptor, has a stronger association with MB leprosy compared to the PB form. As well, the receptor-interacting serine/threonine-protein kinase 2 (*RIPK2*), an adapter protein, leucine repeat region kinase 2 (*LRRK2*) and the multicopperoxidoreductase domain-containing 1 (*LACC1*; previously known as C13orf31) whose putative role in leprosy pathophysiology is still under study, are also known to be more strongly associated with the severe MB form of leprosy[60].

Most of the genes found associated with leprosy *per se* are thought to have a role in innate immunity and bacterial recognition at the early stages of infection while genes associated with leprosy type belong to the acquired immunity arm.[65].

***PARK2/PARCG* locus**

In 2004, an article published in *Nature* reported that variants in an 80-kilobase promoter region shared by both Parkin (*PARK2*) and Parkin co-regulated gene (*PACRG*) were found to be strongly associated with leprosy *per se* in two different populations through large scale linkage and association analysis[66]. The authors discovered, using close to 200 families in the Vietnamese population and almost a thousand unrelated cases and controls in the Brazilian population, that polymorphisms at positions rs1040079 and rs9356058 were risk factors for

leprosy. Indeed, the allelic combination C-T for SNP rs1040079 and rs9356058, respectively, was linked to an increased risk to leprosy compared to the T-C combination for the same SNPs (OR = 5.28; CI 95% = 2.06–13.55).

Despite the striking finding, a group who used the Chinese population for GWAS was unable to repeat the same association between the *PARK2/PACRG* promoter region and leprosy during their replication studies. However, this group was able to include the Parkin protein in a pathway related to *LRRK2* gene which they had replicated in their association study[2]. Very recently, it was confirmed that the linkage disequilibrium pattern and the age-of-onset were critical factors to reproduce association in different populations, and critical role of *PARK2* in leprosy susceptibility was extended to the Indian population[67]. Often the uniqueness of the linkage disequilibrium (LD) pattern in one population can account for the variability of genetic association across populations. Specific statistical and analytical methods have to be applied to overcome this factor.

The SNP rs9356058 that was not found to be associated with leprosy in the Indian population when using a univariate analysis became highly significant when using a multivariate analysis. It was also noticed that stratification for the age at time of diagnosis of the disease might impact on the strength of genetic association to a disease. In fact, excluding the older Vietnamese subjects in the *Nature* publication brought a higher significance to previously identified variants associated with leprosy as well as new SNP discoveries in the *PARK2* region for the disease. As for the SNPrs1040079, despite appropriate statistical analysis and age-of-onset segregation, the variant was not found to be associated to leprosy in

the Indian population. In fact, it is also possible that some genetic associations are specific to an ethnicity. All these findings regarding the *PARK2/PACRG* locus are of great interest, however, a biological cause is yet to be found to explain and confirm the cause of the association to leprosy *per se*. Nevertheless, since *PARK2* is associated with leprosy *per se* (NOT with leprosy type), it is thought that Parkin participates in the innate phase of host defense against *M. leprae*.

It is of great interest to note that genetic associations made with leprosy overlap genetic findings made regarding other important diseases whether they are infectious as tuberculosis[68,69] and typhoid fever[70] or whether they are more chronic and/or complex like Crohn's disease[71] as well as Parkinson's disease[72]. It indicates, of course, common etiologies for similar diseases but also the multiple putative roles of diverse proteins.

Parkin

Parkin and Parkinson's disease

Parkinson's disease is a common disorder of the locomotive system with a prevalence of approximately 1% of individuals over the age of 65[73]. The illness is characterized by tremor at rest, bradykinesia and rigidity[74]. In the later stages of the disease, neuropsychiatric disorders can be noticed as apathy, dementia, irritability or anxiety[73]. The pathology is defined by the reduction of dopaminergic neurons by cellular death and the inclusion of abnormal protein deposits (Lewy bodies) in the substantia nigra located in the midbrain [75].However, the cause of this degenerative neuropathy is still unknown. Despite the ambiguity surrounding the origin of the disease, it is generally recognized that up to 10% of affected individuals tend to have familial history of Parkinson's hence the genetic interest in the study of the disease[7,73].

Named after the disease, Parkin is an E3 ubiquitin ligase encoded by the *PARK2*genein humans. Discovered in the late 1990s, the 1.6 megabase pair gene contains 12 exons (Fig 1.6) mapping to chromosome region 6q25.2-q27 [76].There are at least 7 known isoforms of Parkin, but the isoform 1 is considered to be the canonical version of the protein. Parkin is expressed in abundance in the brain, but it is also expressed a variety of tissues including cardiac, testicular and muscular tissues[77]. It can localize in the nucleus, the endoplasmic reticulum, the mitochondrion[78], but it is mainly a cytoplasmic protein[79].

Mutations in the *PARK2* gene are well known for causing autosomal recessive juvenile Parkinsonism, a form of the disease that normally affects

patients with an average age-of-onset of 30 years old. In fact, Parkin mutations are responsible for half of familial Parkinson's disease cases[80,81]. Moreover, Lewy bodies tend to be absent from brain autopsies of patients with Parkin mutations[82]. There are many hypotheses attempting to explain how the loss of function of Parkin causes the disease, but it is generally thought that the absence of Parkin leads to the accumulation of some neurotoxic proteins possibly causing the death of dopaminergic nerve cells.

Structure

Parkin is a multidomain protein comprising465 amino acids and with a mass of approximately 52 000 Da (Fig. 1.6). Its N-terminal has an ubiquitin-like molecule. The protein contains two really interesting new gene (*RING*) finger motifs (RING1 and RING2) located at its carboxyl end. These motifs are separated by a cystein-rich in-between RINGs region (IBR) which binds to zinc[7]. Recently, a new cystein-rich RING domain (RING0) was identified between the ubiquitin-like domain and the RING1 domain[1]. This characteristics identify Parkin as member of the ring between ring fingers (RBR) protein family[83].

Figure 1.6 Parkin structure
The *PARK2* gene contains 12 exons that codes for a multidomain protein. The UblD is critical for Parkin linking to the proteosome. As for RING domains 1 and 2 they are thought to interact with an E2 ubiquitin-conjugating enzyme and for substrate specificity. IBR and RING0 are important zinc-binding domains[1,7]

Known and putative functions

Parkin is primarily known for its E3 ubiquitin-protein ligase activity. It modulates the ubiquitination process of target proteins by interacting with an ubiquitin-activating enzyme (E1) and an ubiquitin-conjugating enzyme (E2)[7]. The E3 ligase is critical for insuring the substrate specificity at the last step of the covalent conjugation of the molecule[84]. Ubiquitin is a small molecule important for multiple cellular processes like endocytosis, cell cycle control, inflammation and DNA repair. The number of substrates increases at a constant pace[74].

Through its ubiquitin tagging function, Parkin actively participates in the proteosomal degradation pathway of several proteins. In fact, it has been noted that polyubiquitination by Parkin is consistent with an imminent degradation of a substrate, as monoubiquitination indicates another biological regulation[85]. Among Parkin substrates, many are found to be neurotoxic related to Parkinson's pathology such as synphilin-1, glycosylated alpha-synuclein, PAEL-R, cyclin E. Parkin also is capable of auto-ubiquitination leading to its own degradation. As well, the protein can be inactivated through S-nitrolysation which could contribute to juvenile Parkinson's pathology but preventing Parkin to bind to his substrates[86].

Parkin is involved in autophagy regulation. It was reported that Parkin promotes the degradation of impaired and aging mitochondria by relocalization to the organelle. The process is dependent on a serine/threonine kinase, PINK-1 which phosphorylates Parkin's ubiquitin-like domain upon depolarization of the mitochondrial membrane and encourages its recruitment to the proper cellular component which causes an eventual mitophagy[87]. It was also shown that Parkin

inhibits autophagy by stabilizing an anti-apoptotic protein, Bcl-2, through mono-ubiquitination[88]. Additionally, a recent study has demonstrated that Parkin possesses anti-apoptotic and as well as antioxidant functions in neuronal and in myogenic cells[89].

Several studies have hypothesized about a possible tumor suppressor role of Parkin. *PARK2* often shows loss of heterogeneity in tumor biopsies consistent with a function in tumorigenesis[90]. The protein is found to be significantly down-regulated in breast and ovarian cancers[91]. Likewise, *Parkin*-null mice seem to develop hepatocellular carcinomas possibly through the inhibition of caspase which, in return, renders cancerous hepatocytes resistant to apoptosis[92]. The putative role of Parkin as a tumor suppressor is increasingly supported by recent studies although the mechanism is not well understood.

Parkin is an important actor in DNA repair regulation and confers protection against genotoxicity. In fact, DNA damage promotes nuclear translocation of Parkin[93]. Furthermore, genomic material damaged by oxidative stress and UV radiation is restored by physical interaction between Parkin and the proliferation cell nuclear antigen which modulates DNA excision repair[85]. Additionally, Parkin helps to maintain mitochondria genomic integrity by promoting mitochondria DNA repair and increasing the replication of the mitochondrial genome[94].

Parkin in infection and immunity

There is already strong evidence for genetic association between the *PARK2* locus and some infectious diseases. However, a biological and molecular basis explaining the relationship between the Parkin protein and immune processes is still lacking. In the case of leprosy, the expression of Parkin in cell types and tissues of interest, namely Schwann cells and monocytes/macrophages, reaffirms the relationship of Parkin with immunological pathways [66].

It has been suggested by Schurr et al. (2006) that Parkin could interact with the host response with *M. leprae*[8]. It was recognized that nitric oxide (NO) and reactive oxygen intermediates (ROI) not only can by produced by Schwann cells and inflammatory macrophages following nerve damage, but also induce apoptosis in Schwann cells. Furthermore, apoptosis negatively impacts *M. leprae* survival. Since Parkin is known to act as an anti-oxidant in the Drosophila model, the group suggested that susceptible or resistant alleles located in the *PARK2* promoter could result in differential expression of Parkin causing dissimilar host defense against *M. leprae*. One element to mention is not only has Parkin been shown to be an oxidative stress regulator, but there are also signs showing that oxidative stress alters Parkin's solubility leading to its degradation[95]. Hence, Parkin down-regulation could lead to host susceptibility to *M. leprae*.

On the other hand, another pathway was described to potentially explain the relationship between Parkin and immunity. It was proposed that the protein is regulated by the nuclear factor kappa-B (NF-κB) pathway. NF-κB is an important player in immune responses, and could repress *PARK2* transcription[5]. In fact,

Parkin levels could possibly regulate the expression of known inflammatory genes such as tumor necrosis factor (*TNF*). It was proposed that Parkin in its natural, unstimulated state, limits the expression of inflammatory factors. Through infection, NF-κB possibly represses Parkin to allow the proper inflammatory elements to be expressed, suggesting that Parkin's down-regulation is necessary for the innate immune response[5].

PACRG

PACRG is a gene spanning over 0.6 Mb located head to head and antisense to the Parkin gene. The two, share a promoter. The gene is often co-expressed with *PARK2*. Little is still known about the function of the gene[96]. However, it is currently accepted that *PACRG* acts as part of a chaperone associated to heat shock proteins 70 and 90 as well as chaperonin components[97].

Mouse model

The intensive work that has been done in genomics in last decade has given scientists new information on how to better study human diseases using animal models. Although only 40% of the *Mus musculus* genomic sequence aligns with the *Homo sapiens* sequence, their genomes share 90% conserved synteny. The mouse genome is approximately 14% less than the human one probably due to a higher rate of deletion in the mouse, but both species possess a similar number of protein-coding genes. Most importantly, the number of genes without homology in one genus or the other is less than a percent[98].

The mouse has become the most common animal model to study human diseases due to multiple factors. The important genomic, proteomic, biologic and functional similarities between the two species cannot be ignored. Most homologs have the same function in both species and common functional pathways are the rule more than the exception. These homologies account for strong conservation of molecular pathways in health, but also in disease-state[99]. Moreover, mice area useful complement to human studies due to their fast generation time, relatively low maintenance costs and short lifespan. Also, mice are easy to handle, and can be manipulated and studied at the molecular level in a straightforward manner[100].

Transgenic mice: The *Hprt* locus

The capacity to modify gene sequences in different cell types or organisms has permitted to us better define their role in a biological system. Moreover, the ability to reproduce human phenotypes in other species has contributed to the understanding of a great number of diseases. Especially mouse transgenesis has

been found to be extremely useful to trace mechanisms of pathogenesis. However, the modification of genomic content is a challenging and complex procedure. Some challenges encountered in transgenesis with various methods are incorrect gene dosage (more than one transgene inserted in the genome), unknown locus of insertion, and disruption of endogenous genes.

In the mid-1990s, a scientific group from North Carolina described a method to insert in a directed fashion, a single copy transgene in mouse germ lines using homologous recombination in embryonic stem cells[98,101]. This was achieved using the hypoxanthine phosphoribosyltransferase (*Hprt*) locus located on the X chromosome as target for transgene insertion (Fig. 1.7). The choice of *Hprt* as the insertion location was done to insure confident and reliable results since *Hprt* is a housekeeping gene expressed in virtually every cell type at all stages of development. Therefore, the gene lies in a constitutively open chromatin region favorable for transcription. The genome location assures that promoters, enhancers, and repressors maintain their properties and that the transgene they influence maintains its transcription pattern[101].

Once the homologous recombination has occurred inserting the transgene at the 5'end of the Hprt gene, modified embryonic stem cells can then be injected in murine blastocysts which will be inserted in a pseudopregnant female. Offspring will be chimeras as the introduced cells will develop in various organs of the offspring. The tissue mosaicism in these chimeras will result in some males to produce sperms containing the transgene. A variable proportion of offspring obtained from chimeric fathers will have been fertilized by transgenic sperm and

result in fully transgenic mice. These mice will be selected for further breeding and through various crosses transgenic mouse strains with different genetic backgrounds can be obtained for experimental studies[101].

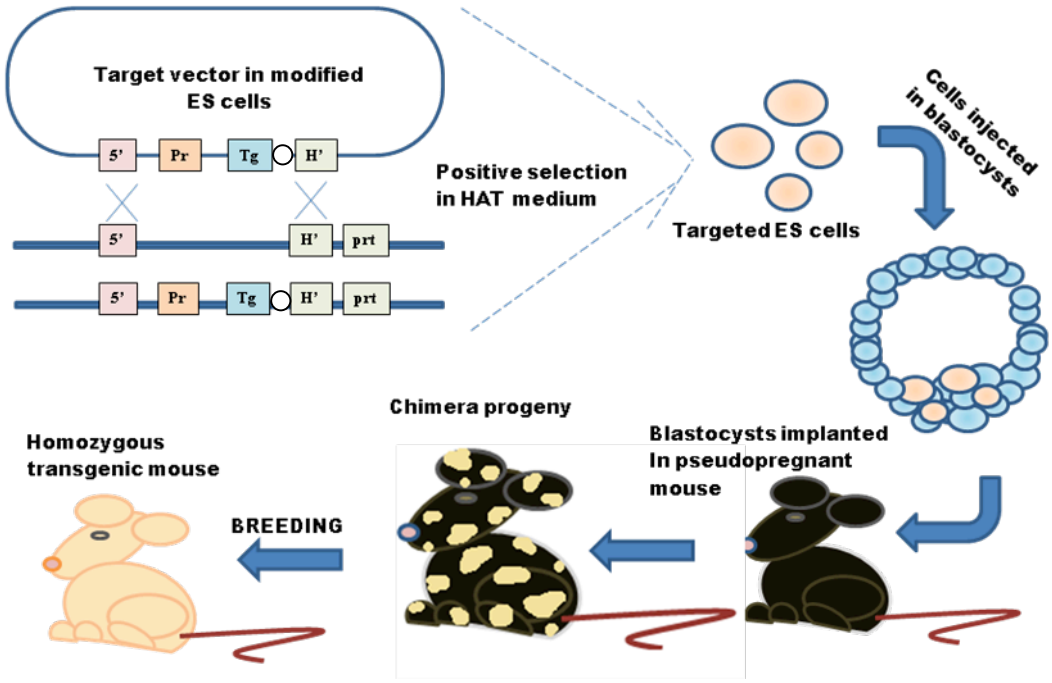

Figure 1.7 Mouse transgenesis using the *Hprt* locus
The *Hprt* locus located on the X chromosome is useful for mouse transgenesis. First, a target vector is engineered with homology in the region 5' of the *Hprt* gene (5'), a promoter (Pr) and the transgene of interest (Tr). BPES cells, embryonic stem cells lacking a functional element of *Hprt* (○) are used. The BPES cells are transfected with the target vector and through chromosomal recombination the promoter and transgene are inserted in the genome a specific locus and *Hprt*'s functionality is reestablished. The cells are incubated with hypoxanthine-aminopterin-thymidine (HAT) medium were only successfully targeted cells will survive. These cells will then be injected in mouse blastocysts that will in return be implanted in a pseudopregnant mouse. The foster mouse will give birth to chimeras as the targeted cells will have been inserted in various tissues of the mouse. Through a series of breeding and backcrosses one can obtain a pure transgenic strain.

Mouse genetics for human genetics

Mouse genetics traces back to the beginning of human civilization. The humans back then, were able to observe different colour-coated mice and even kept records of their discoveries[98]. By the 18th century, mouse domestication was current in Asia and Europeans had started to import some strains for inbreeding. The scientific study in mouse genetics became relevant in the 20th century when the community started revisiting Mendel's law of inheritance and noticing the multiple variations across mouse strains. From then, strain engineering became a serious research matter and resulted in the generation of the various laboratory mouse strains used today[102].

Advances in genomics and in DNA technologies applied to the mouse permitted the discovery of many polymorphisms responsible for strain differences. Tools such as PCR were fundamental for the genotyping of restriction-fragment length polymorphisms and microsatellites[98]. Today, thanks to next-generation sequencing, there are close to 20 genome sequences of commonly used laboratory mouse strains publicly available, giving us important information on the phylogenic tree of these mice as well as functional indications on allelic variations[103]. Using various mouse strains can to some extent, reproduce the genetic variability across populations when studying genetic traits[104].

Currently, there are 1,134 human diseases for which mouse models exist[105]. The International Knock Mouse Consortium actively contributes to the increase of this number by aiming to create a knock out strain from every gene in the mouse genome[106]. This approach will highly contribute to the study of

monogenic genetic diseases, but also help understand the contribution of specific genes to complex diseases. Mouse genetics has been proven to be extremely useful in the study human genetics applied to infectious diseases. To use a specific example, the discovery of the Solute Carrier Family 11, member 1 (Slc11a1)[107,108] highly contributed to the understanding of human tuberculosis pathogenesis through mouse genetic studies[109].

Study of Parkin in mice

Park2, located on chromosome 17 of the mouse is quite conserved when compared to the human *PARK2* gene[110]. *Parkin*-null mice have been engineered in an attempt to reproduce the same phenotype found in Parkinson's patients. However, the results are inconsistent and the loss of dopaminergic neurons is not a phenotype of these mice. Also, the clinical manifestations displayed by individuals afflicted by Parkinsonism were not yet successfully reproduced in the mouse when knocking out Parkin[111]. Despite these facts, the study of Parkin in the mouse is still relevant for understanding the function of the protein at the molecular level.

CHAPTER 2

Experimental strategy

In order the measure the levels of expression of human *PARK2*, two strains of transgenic mice had been engineered. The first strain, coined R3, carries a promoter construct with the allele combination associated with resistance to leprosy. The second strain, coined S5, carries the alleles associated with leprosy susceptibility[66]. These transgenic mice allow to estimate the activity of the two human *PARK2* promoter constructs under different conditions through measurement of luciferase reporter activity. Hence, *PARK2* promoter activity is well reflected by the amount of luminescence produced during the conversion ofluciferin to oxyluceferin, a reaction catalyzed by the luciferase enzyme.

The experimental approach pursued in my study consisted of infecting both strains of transgenic mice and harvesting organs of interest at specific time points according to the pathogen used (Fig. 2.1). A part of the organ collected was set aside for RNA extraction to measure endogenous mouse *Park2*. This was done to compare in control experiments human *PARK2* expression from the transgene with endogenous murine *Park2* expression. The remaining organ was used to quantify luciferase activity. To assure an accurate estimate of luciferase activity, the luminescence measured was normalized to total lysate protein. *PARK2* expression was compared between mice from both strains under normal conditions and after infection with different pathogens. This allowed to assess the impact of pathogens on human *PARK2* promoter activity in general and to

determine a possible modulatory role of alleles associated with resistance and susceptibility to leprosy on promoter activity.

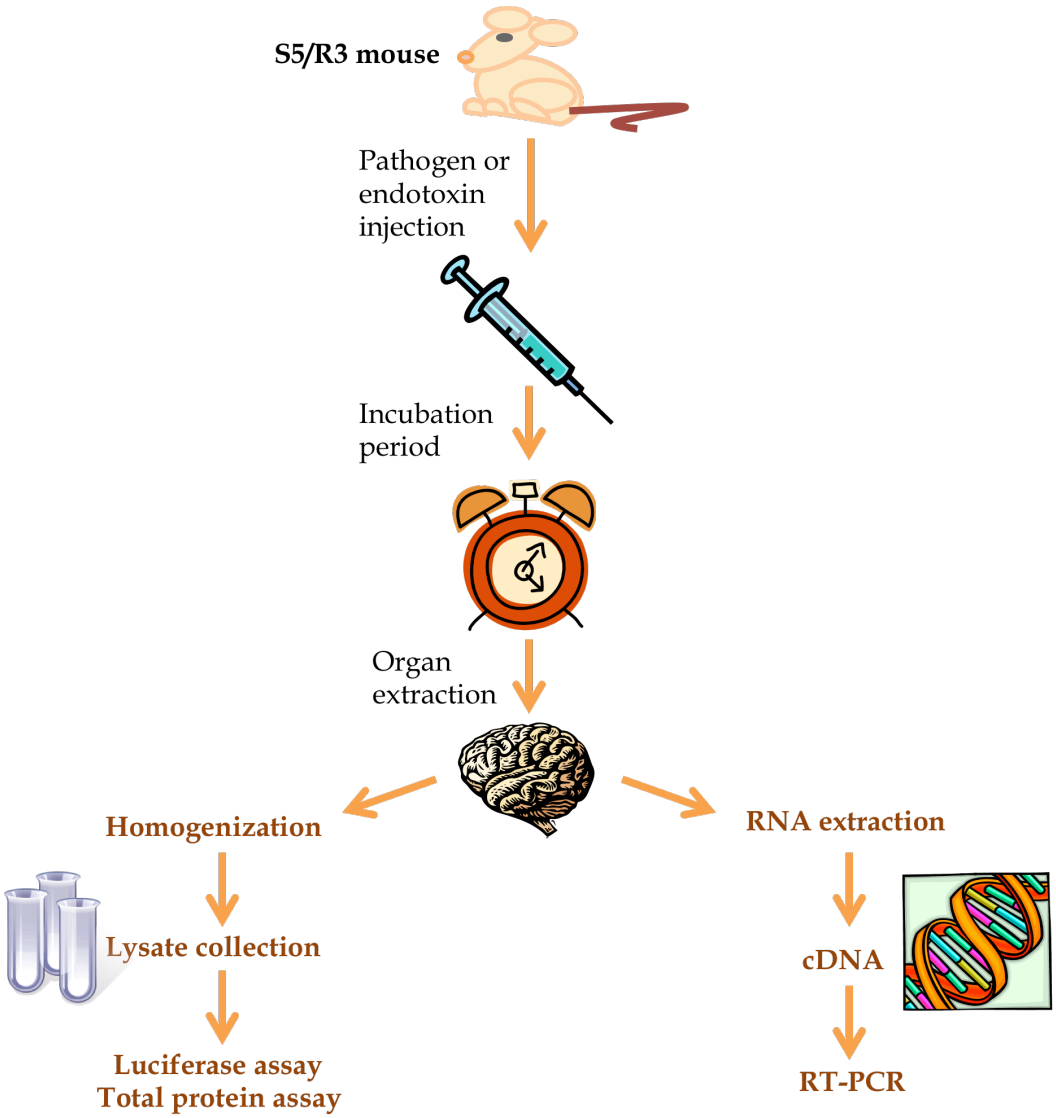

Figure 2.1 Flowchart of experimental approach.
Pathogenic agents are injected IV or IP (depending on the agent) in mice of both strains (R3 and S5). At various time points, mice are sacrificed and organs of interest are collected. A portion of each organ is put aside in RNA stabilizing solution until RNA extraction for reverse transcription in complementary DNA (cDNA) and subsequent real-time polymerase chain reaction (RT-PCR). The remainder of the organ is homogenized and the lysates are collected for luciferase and whole protein assays.

The polymorphisms rs9356058 and rs1040079 located in the shared *PARK2/PACRG* promoter region had been identified as strong leprosy susceptibility factors[66]. However, SNP rs1040079 is 62,718 bp telomeric to SNP rs9356058 which precluded use of both SNPs in the same expression construct. On the other hand, rs1040079 is part of a group of highly correlated SNPs, a so-called SNP bin, that extends from rs1040079 all the way to the vicinity of rs9356058. All members of this SNP bin show strong evidence for association with leprosy and, thus, are leprosy risk factors. One of those SNPs is rs2276201 which relative to rs9356058 is 1,812 bp closer to *PARK2*. The SNP is located in the core promoter region at 663bp upstream of the transcriptional start site of *PARK2*. In subsequent experiments, rs2276201 was shown to be an independent risk factor for leprosy in both the Vietnamese and the North Indian population (Fava et al Hum Genet 2013). Hence, the promoter construct used for the transgenesis contained the resistance and susceptibility configurations of SNP rs2276201 and rs9356058 (Fig. 2.2). The R3 and S5 strains carry the *Hprt* (p)*PARK2* + *FLuc* construct on a C57BL/6J background (cf methods).

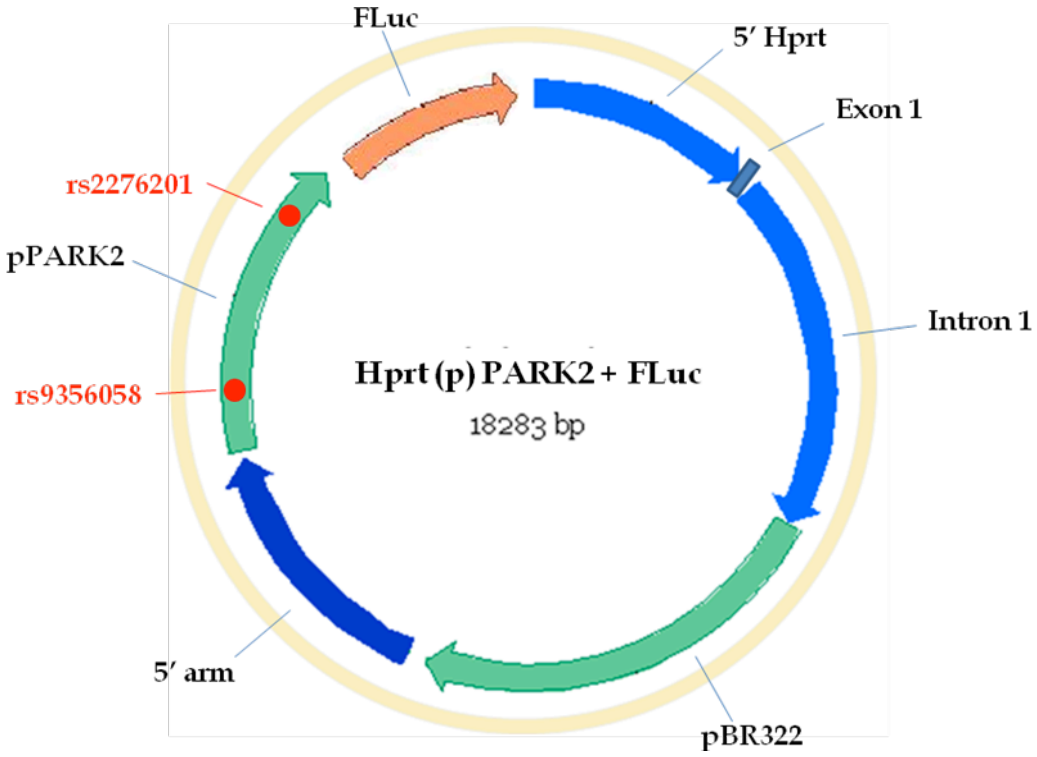

Figure 2.2 Schematic of target constructs used for transgenesis
The *Hprt* (p)*PARK2* + *FLuc* constructs originated from the pBR322 bacterial plasmid. A 5' homology arm (5'arm) as well as the 5'flanking region, exon 1 and intron 1 for Hprt target the construct by homologous recombination to the X chromosome *Hprt* gene. In this way, only one copy of the transgene is integrated per genome. The human *PARK2* promoter region (pPARK2) with SNPs rs2276201 and rs9356058 in the "resistant" and "susceptible" configuration is included. Finally, firefly luciferase (*FLuc*) is added to allow quantitation of *PARK2* promoter activity. In total, the constructs are 18,283 bp in size.

CHAPTER 3

Results

Expression of the pPARK2/FLuc transgenes in mouse organs

In a first set of experiments we asked if any expression of the transgene could be detected, and if so, how mouse organs differed in their level of *PARK2* promoter activity. Hence, we determined the constitutive level of luciferase reporter activity in brain, lung and spleen of both transgenic strains.

Luciferase activity could be detected in all three organs from both R3 and S5 mice. Moreover, luciferase activity was detected at variable levels across organs suggesting organ-specific expression of the *PARK2* promoter (Fig. 3.1). We observed that constitutive expression of *PARK2* was highest in the brain, followed by the spleen and finally the lung in mice from both R3 and S5 strains. These results were consistent with the relative expression levels of the *PARK2* gene in human brain, spleen and lung, and suggested that the transgenic constructs used in our experiments faithfully reflected the promoter activity of the human gene. Finally, we compared expression of the constructs in organs of R3 and S5 mice. In healthy, uninfected mice, constitutive luciferase activity across all three organs was similar for both strains with no statistical difference between constructs (Fig. 3.1).

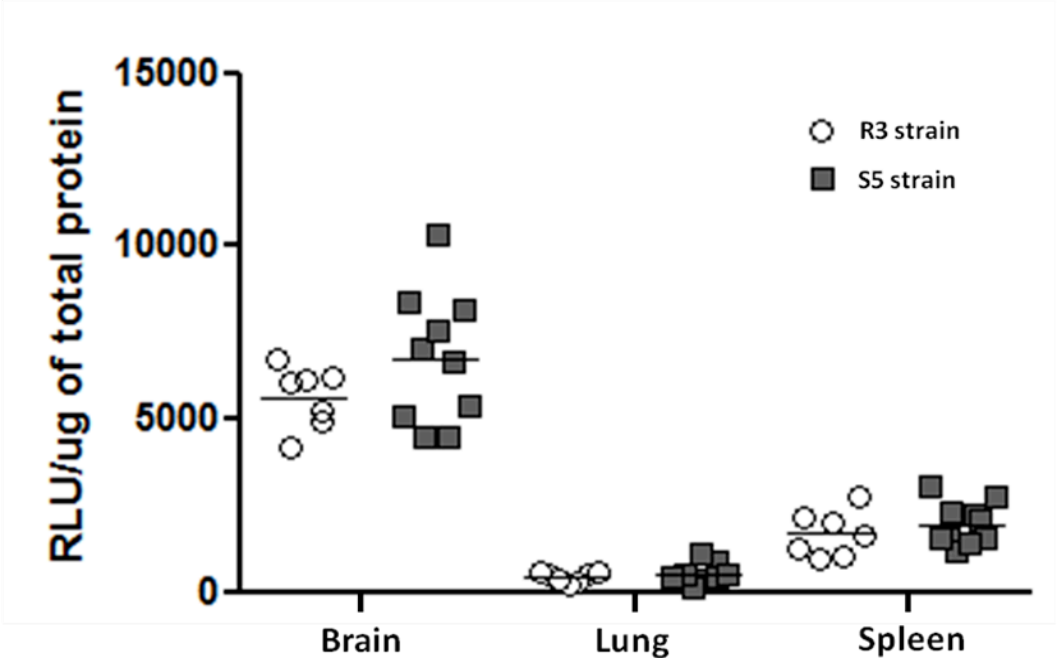

Figure 3.1 Constitutive pPARK/FLuc expression in mouse follows the same trend as *PARK2* expression in human tissue
Whole organ lysates from brain, lung and spleen were collected from R3 and S5 mice. Each symbol represents one mouse. The results are displayed in RLU (relative light unit) for luciferase quantification per μg of total protein. The bars represent the mean value per group.

pPARK2/ FLuc transgenes are down-regulated by LPS and intracellular infection

Since constitutive expression of the pPARK2/FLuc did not differ between R3 and S5, we investigated the impact of infection on the reporter activity. Since M. leprae does not grow in normal mice we chose Bacille Calmette-Guérin (BCG), strain Russia, as infectious pathogen. BCG strains are attenuated forms of *Mycobacterium bovis* that are being widely used as vaccines against tuberculosis [112]. BCG Russia is the only BCG strains that grows to high titers in the lungs and spleens of infected mice [104].

Mice from both strains were infected with the same dose of BCG Russia by intravenous (IV) injection. At 5 weeks post infection, the infected mice as well as controls for each strain were sacrificed. Organs (brain, lung and spleen) were collected and processed for luciferase quantification. When proceeding with the organ extraction, splenomegaly and often lung granulomas were visible in infected mice and provided independent proof of ongoing infections in these mice. Luciferase assays were conducted for all organs and results were normalized to total protein concentration in lysates.

Surprisingly, for mice from both strains the level of luciferase activity per μg of protein was significantly lower in the spleen of infected animals as compared to controls (Fig. 3.2). These results suggested a repression of *PARK2* promoter activity in spleens upon infection with BCG. There was no such repression of reporter activity in the lung (Fig. 3.3) or the brain (Fig. 3.4) following infection with BCG. The luciferase activity was repressed in mice of both sexes. In male mice, the difference in expression between infected and control mice was significant for both R3 ($p = 0.0159$) and S5 mice ($p = 0.0217$) (Fig. 3.2A). Compared to males, the suppression of *PARK2* by BCG was stronger in females of both R3 ($p = 0.0004$) and S5 mice ($p = 0.0001$) (Fig. 3.2B). There was no significant difference in extent of down-regulation of luciferase activity following BCG infection between the two strains in any of the three organs studied. These results suggested that BCG infection did not trigger *PARK2* down-regulation in a leprosy susceptibility allele specific fashion. The experiment was repeated 3 times with similar results.

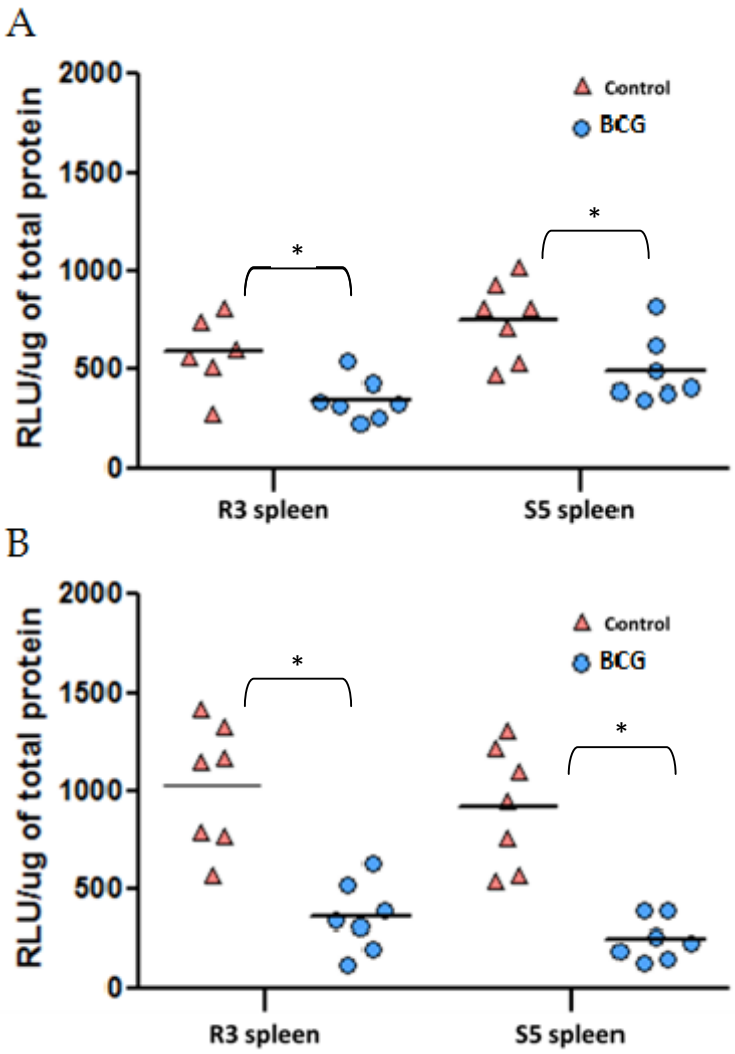

Figure 3.2 BCG down-regulates the pPARK2/FLuc expression in the spleen

Mice from both R3 and S5 strains received 100µL of PBS diluted BCG through IV injection. Infected and control mice were euthanized 5 weeks post-injection by CO_2 asphyxiation. The organs were extracted and processed for luciferase activity (RLU) and BCA assay (µg of total protein). The ratio is plotted on the graph above. (A)Male (B)Female. Unpaired t-tests were used to determine significance (*) $P < 0.05$. The bars represent the mean value per group.

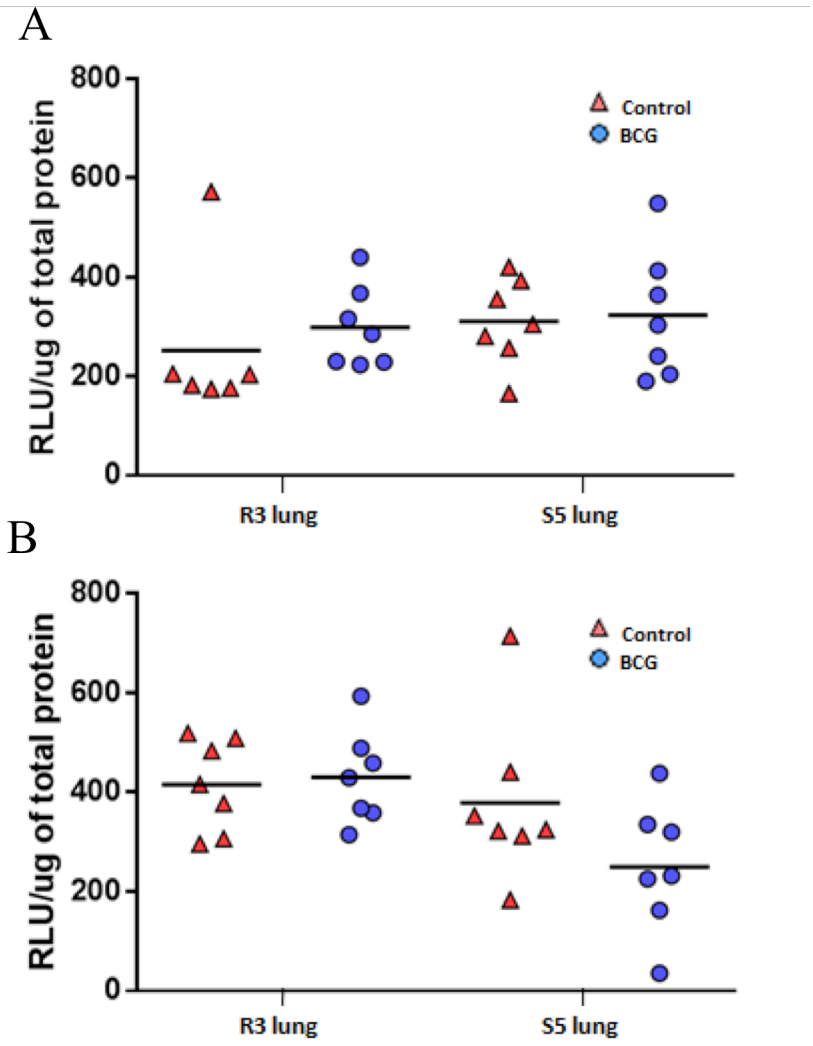

Figure 3.3 BCG infection has not impact on pPARK2/FLuc expression in the lung

Mice from both R3 and S5 strains received 100µL of PBS diluted BCG through IV injection. Infected and control mice were euthanized 5 weeks post-injection by CO_2 asphyxiation. The organs were extracted and processed for luciferase activity (RLU) and BCA assay (µg of total protein). The ratio is plotted on the graph above. (A)Male (B)Female. Unpaired t-tests were used to determine significance. The bars represent the mean value per group.

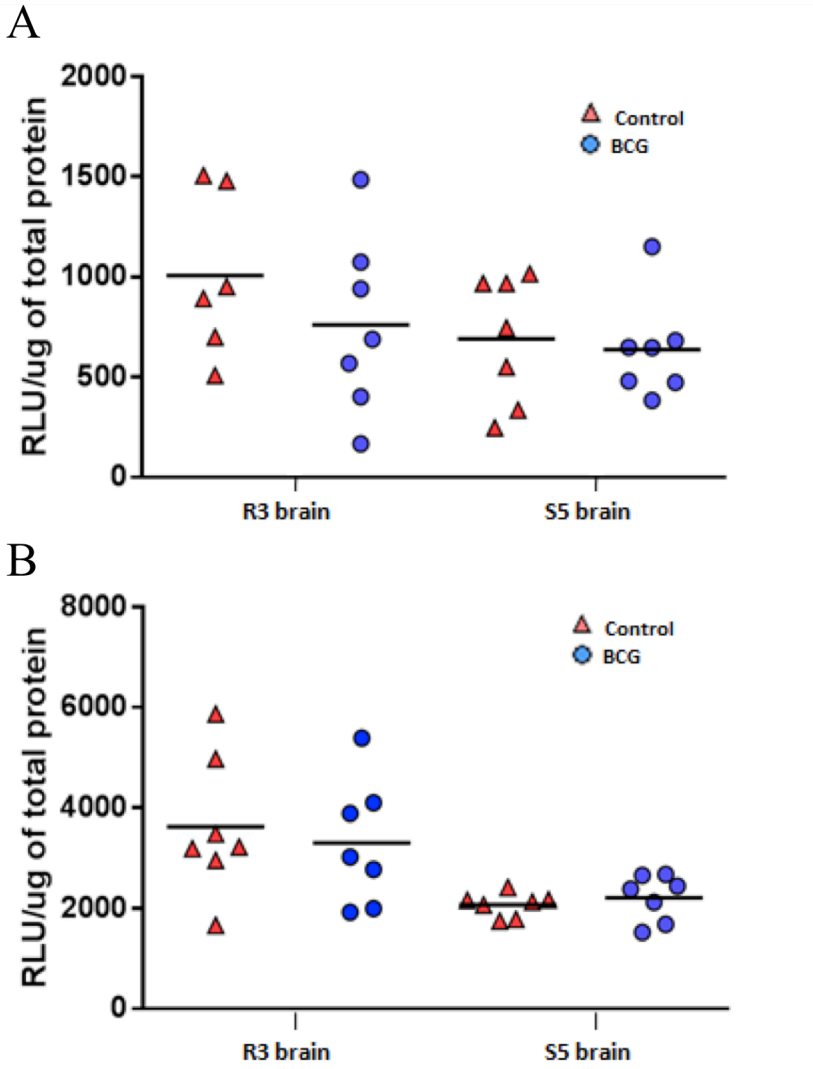

Figure 3.4 BCG infection has not impact on pPARK2/FLuc expression in the brain

Mice from both R3 and S5 strains received 100 µl of PBS diluted BCG through IV injection. Infected and control mice were euthanized 5 weeks post-injection by CO_2 asphyxiation. The organs were extracted and processed for luciferase activity (RLU) and BCA assay (µg of total protein). The ratio is plotted on the graph above. (A)Male (B)Female. Unpaired t-tests were used to determine significance. The bars represent the mean value per group.

Next we asked if the down-regulation of pPARK2/FLuc expression was BCG specific. For this purpose, we chose the gram-negative fast-growing bacterium, *Salmonella Typhimurium*. This pathogen is closely related to *S.Typhi* and *S. Paratyphi* responsible for causing typhoid and paratyphoid fever. Actually, the SNPs rs9356058 and rs2276201 included in our constructs were found to be risk factors of typhoid disease[70]. Moreover, mice on a C56BL/6 background like our two transgenic strains are highly susceptible to the *S. Typhimurium*.

Mice of both strains and both sexes were injected with 1000 CFU *S. Typhimurium* through IV injection. Three days following IV infection, the mice and their controls were sacrificed and organs were processed as described for the BCG experiment.

We observed a repression of luciferase activity in splenic tissue after infection (Fig. 3.5). There were no reproducible differences in reporter activity in brain or lung after infection. Similarly, the R3 and the S5 constructs were similarly affected by the infection with no consistent and/or significant differences between the two constructs. However, the down-regulation was more pronounced in male mice ($p = 0.0015$ for R3 and $p < 0.0001$ for S5) (Fig. 3.5A) as compared to females ($p = 0.0059$ for R3 and $p= 0.0001$) (Fig. 3.5B). The reason for the opposite gender effects in BCG and *S. Typhimurium* infection are not known. Nevertheless, these results clearly demonstrated that down-regulation of *PARK2* promoter activity following infection with intracellular pathogens was not mycobacteria-specific.

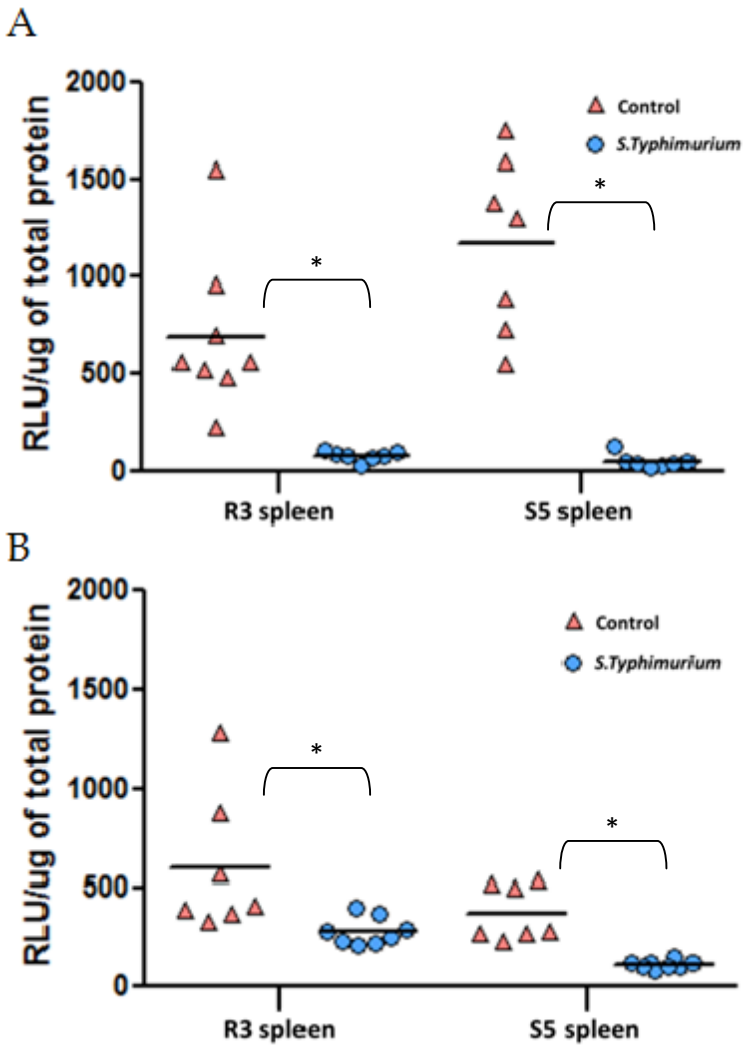

Figure 3.5 *S. Typhimurium* represses the pPARK2/FLuc expression in the spleen

Mice from both R3 and S5 strains received 1000 CFU of *S. Typhimurium* through IV injection. Infected and control mice were euthanized 3 days post-injection by CO_2 asphyxiation. The organs were extracted and processed for luciferase and BCA assay. (A)Male (B)Female. Unpaired t-tests are used to determine significant difference (*) $P < 0.05$. The bars represent the mean value per group.

After finding similar pattern of repression of human *PARK2* promoter activity following infection with BCG, a mycobacterium and *S. Typhimurium*, a gram-negative enterobacterium, we were interested to know if endotoxin would trigger the same response as the two pathogens. In fact, a previous study had shown down-regulation of Parkin in primary murine microglia after lipoposaccharide (LPS) exposure[5]. We sought to investigate the effect of LPS in whole tissue lysates. Since the microglia findings were done *in vitro*, it was relevant to determine the outcome in an *in vivo* system.

We started with intraperitoneal (IP) injections of a phosphate buffer solution (PBS) dilution of 100 μg of LPS from the *E. coli* strain O55:B5 in R3 and S5 mice from both sexes and repeated the protocol employed in the BCG and *S. Typhimurium* experiments. The pPARK2/FLuc expression results were statistically analyzed through.

Surprisingly, we did not observe a repression of the human *PARK2* promoter in whole brain tissue after LPS exposure which does not follow the findings in Tran et al., 2011 (data not shown). Although we were not able to observe a similar repression of the *PARK2* promoter in the brain, we were able to observe statistically convincing repression of the promoter in the spleen of female mice of R3 ($p = 0.0038$) and S5 mice ($p = 0.0084$), respectively (Fig. 3.6B). We failed to observe down-regulation of the luciferase activity in male mice from R3 ($p = 0.6976$) S5 strains ($p = 0.3082$) (Fig. 3.6A). Likewise, we detected no

difference of promoter activities ($p > 0.1$) in lung tissues of both sexes (data not shown).

We performed a kinetic study of LPS induced down-regulation of luciferase activity down-regulation with 2 additional time points (12 h and 24h). In these experiments, we injected 100 μg of a very potent LPS from *E. Coli* strain K-275.There was no statistical difference in human *PARK2* promoter activity between control and infected mice for all three organs at the 12-hour time point (data not shown). However, at the 24-hour time-point we observed a trend for repression ($p = 0.1341$ for R3 and $p = 0.0215$ for S5) of the reporters in the spleen of female mice. This finding suggested a sex and time sensitive modulation of the *PARK2* promoter regulation following exposure to LPS. A major aspect of our experiments was the study of expression differences between the two reporter constructs whether at constitutive state or after immune stimulation by bacterial infection or endotoxin challenge. At times, we observed a striking difference between the R3 strain and the S5 strain in terms of luciferase expression in individual experiments (data not shown). However, these results were not consistently reproducible (data not shown). While there was a tendency for higher expression of the reporter in the S5 mouse strain, it was not possible to observe this difference reproducibly. Despite the fact that using the pPARK2/Luc reporter model is very convenient since it is easily detectable, the reporters were not reliable enough in this respect to clearly give an indication on the impact of the SNP combination in the construct on *PARK2* promoter activity.

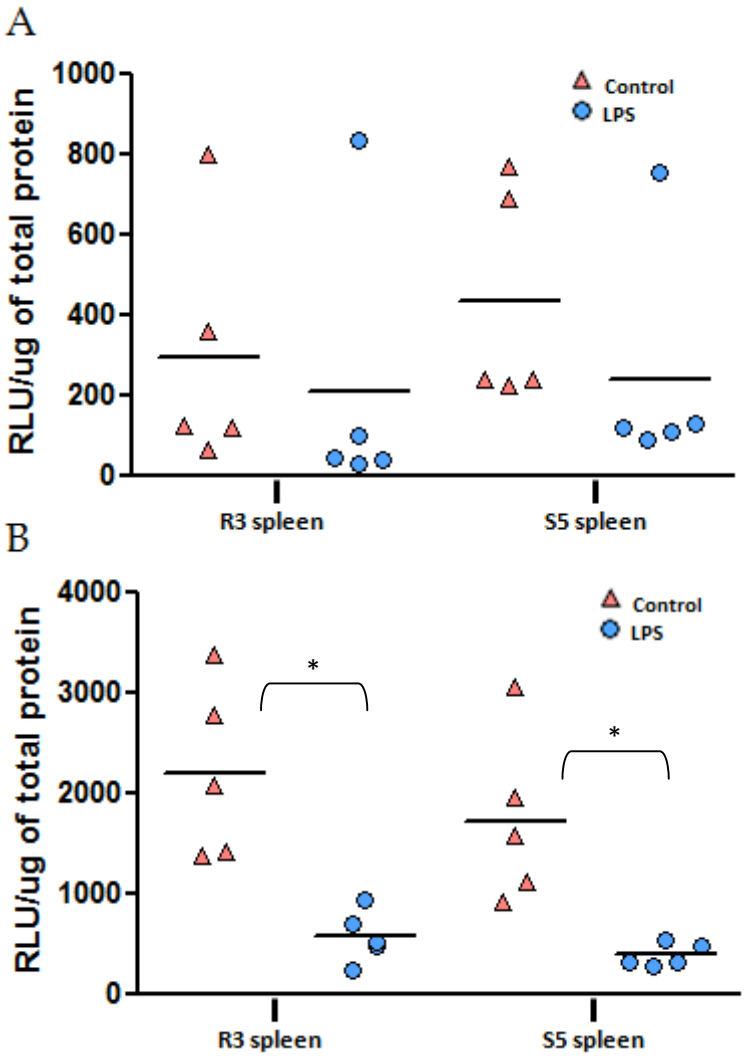

Figure 3.6 LPS down-regulates the pPARK2/FLuc expression in the spleen

Mice from both R3 and S5 strains received 100 µg of LPS (*E. coli* strain O55:05) in 0.5mL PBS through IP injection. Infected and control mice were euthanized 48 hours post-injection. The organs were extracted and processed for luciferase and BCA assays. (A)Male (B)Female. Unpaired t-tests are used to determine significant difference (*) $P < 0.05$. The bars represent the mean value per group.

pPARK2/FLuc transgenes display poor correlation with endogenous Park2

Having examined the impact of infection on the human *PARK2* regulation in the mouse system, we tested the effect of the same pathogens on endogenous levels of murine *Park2*. To investigate the question, in each experiment that assayed transgene expression a portion of organ had been preserved for RNA extraction. RNA was extracted from brain and spleen and for both strains for some male and female mice. Technically, it was not expected to observe a strain difference since the constructs did not affect endogenous *Park2*. cDNA for each sample was made through reverse transcription and used for real-time PCR (RT-PCR). *Gapdh* was chosen as housekeeping gene for normalization of gene expression. First, we analyzed the correlation of reporter construct activity with endogenous mouse Park2 levels in uninfected control mice. We detected no correlation between both alleles of the transgene and the corresponding endogenous Park2 expression. This was true for brain and spleen, as well as males and females. We also found absence of correlation in BCG, Salmonella and LPS stimulated mice as well as in control mice (Table 3.1). In fact, all correlation coefficients (R^2) were far below 0.5 a generally accepted cut-off for moderate correlation

Organ	Pathogen (control and infected mice)	Sex	R^2 with pPARK2/FLuc
Brain	BCG	Male	0.00309
Brain	LPS	Male	0.00057
Brain	LPS	Female	0.0167
Brain	*Salmonella*	Male	0.15699
Brain	*Salmonella*	Female	0.11881
Spleen	LPS	Male	0.001
Spleen	LPS	Female	0.42209
Spleen	*Salmonella*	Male	0.22883

Table 3.1: pPARK2/FLuc expression does not correlate RNA expression
R^2 = coefficient of determination

Endogenous Park2 expression is repressed in the spleen by LPS and intracellular infection

Importantly, bacterial infection and endotoxin exposure caused down-regulation of mouse *Park2* in spleen for both males and females (Fig. 3.7). This finding is consistent with the repression of the *PARK2* promoter construct following BCG and *S. Typhimurium* infection. We detected small fold differences in *Park2* levels between control and BCG infected mice from R3 (-2.27) and S5 strains (-3.40) (Fig. 3.7A). While the down-regulation of *Park2* levels following BCG infection in mouse spleen was not statistically significant there was trend for repression of *Park2*. A very pronounced down-regulation of *Park2* was observed in the spleens of male mice from R3 (-1027) and S5 (-2703) strains following infection with *S. Typhimurium* (Fig. 3.7C). Interestingly, LPS exposure down-regulated Park2 in males of mice from both strains (-48 for R3 and -358 for S5)

(Fig. 3.7B). The same exposure induced a significantly greater fold differences in females (-6099 for R3 and - 1671 for S5) (Fig 3.9A). This result suggested that while the extent of *Park2* regulation differed significantly between the human promoter construct the endogenous mouse gene, the response to inflammatory insult was similar. This suggests that extent gene expression was regulated by factor binding sites outside of the human construct or that different transcription factors regulated extent of gene expression in both species.

The impact of immune stimulation on endogenous *Park2* expression was also studied in mouse brain tissue. As observed with the pPARK2/FLuc constructs, Park2 expression in the brain did not vary upon stimulation with any of the agents (Fig. 3.8 and Fig 3.9B). BCG and *S. Typhimurium*, are bacterial pathogens and do not cross the brain-blood barrier. Consequently, it was expected to see no modulation of *Park2* expression after infection. However, LPS, a fat-soluble molecule, is known to reach and penetrate neurons and lack of effect on *Park2* brain levels is not clear.

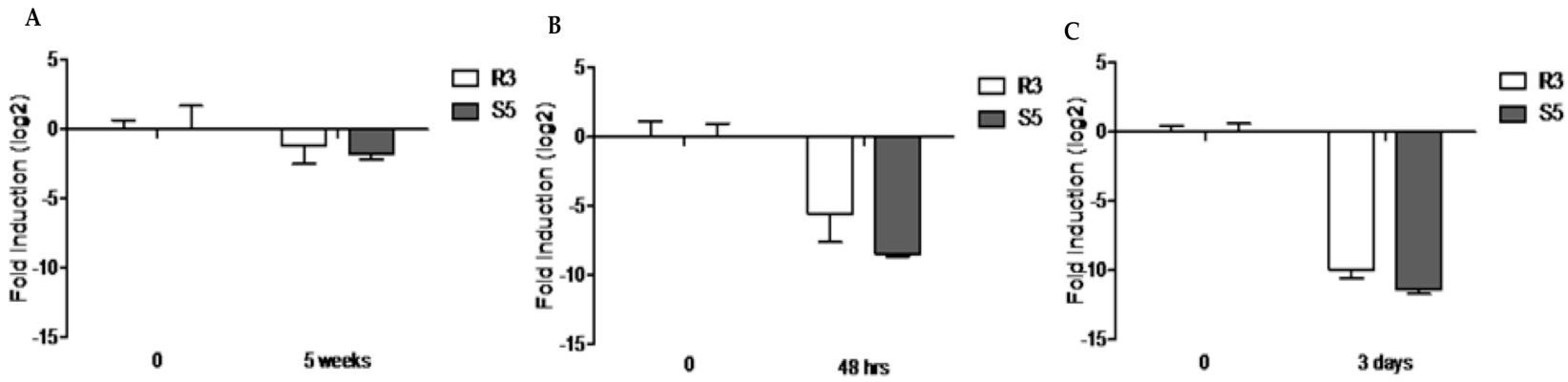

Figure 3.7 LPS challenge, BCG and *S. Typhimurium* infection down-regulate endogenous *Park2* expression in the spleen

RT-PCR for murine endogenous *Park2* in male spleens. *Gapdh* was used as the housekeeping gene. Results for RT-PCR are presented as fold induction. For BCG infection, 3 mice per group were used (A), for LPS injection 5 mice per group were used (B), or the *S. Typhimurium* experiment (C), 3 mice per group in both infected and controls were used for the R3 strain and n=4 per group for the S5 strain.

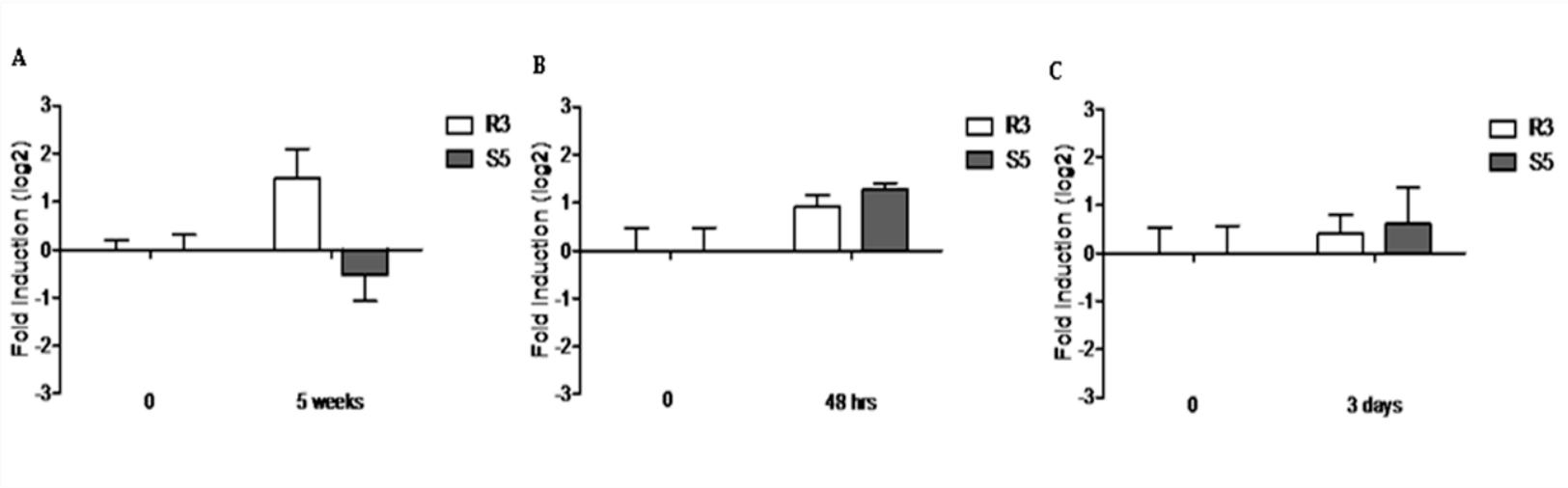

Figure 3.8 LPS challenge, BCG and *S. Typhimurium* infection have no statistical effect on *Park2* expression in murine whole brain tissue

RT-PCR for murine endogenous *Park2* in male brains. *Gapdh* was used as the housekeeping gene. Results for RT-PCR are presented as fold induction. For BCG infection, 3 mice per group were used (A), for LPS injection 5 mice per group were used (B), or the *S. Typhimurium* experiment (C), 3 mice per group in both infected and controls were used for the R3 strain and n=4 per group for the S5 strain.

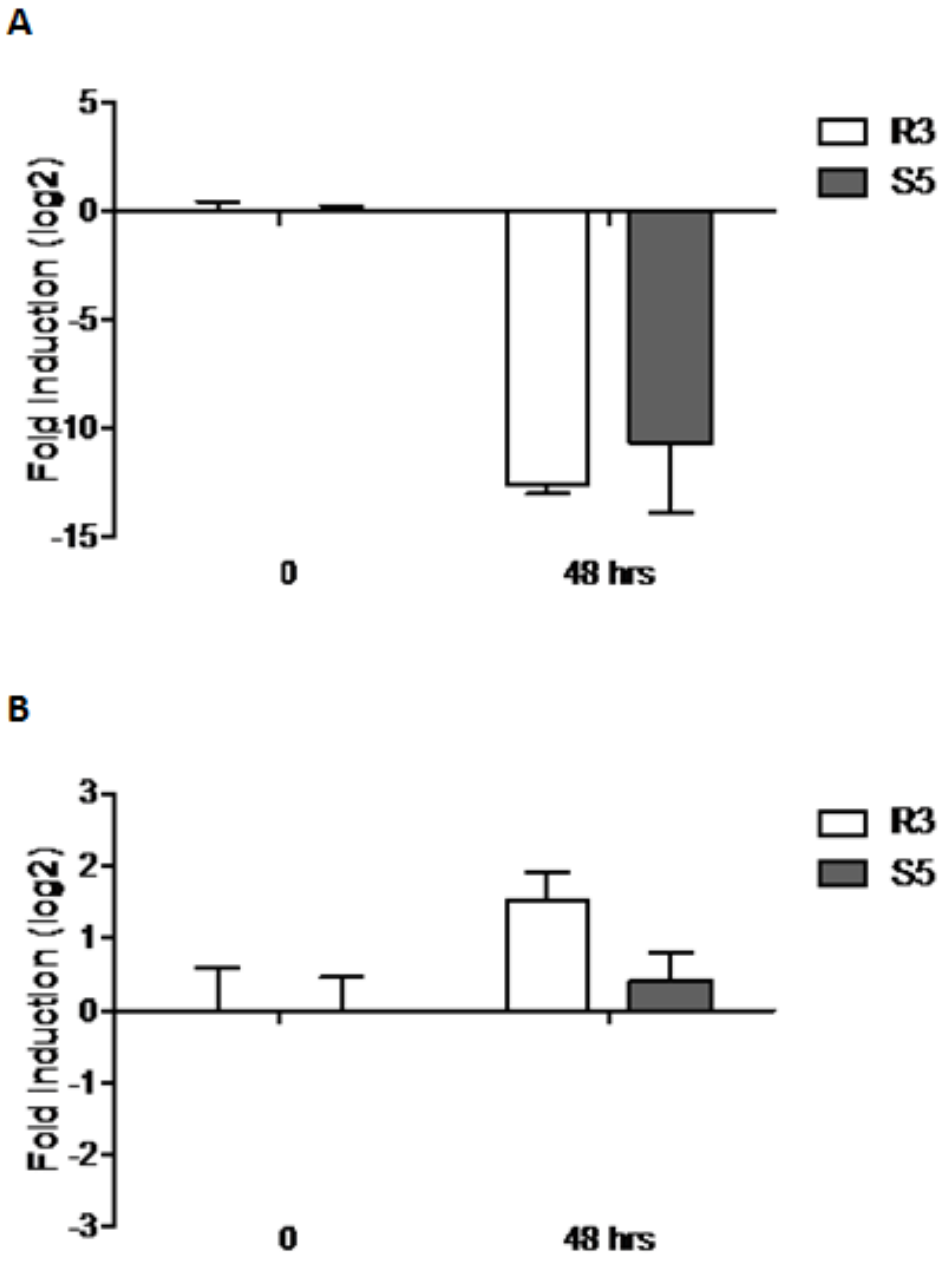

Figure 3.9 LPS challenge has a strong effect on *Park2* expression in female murine spleen
RT-PCR for murine endogenous *Park2* in female spleens (A) and brains (B). *Gapdh* was used as the housekeeping gene. Results for RT-PCR are presented as fold induction. 5 mice per group were used.

CHAPTER 4

Discussion

We showed here that a human *PARK2* promoter constructs in transgenic mice displayed promoter activities in brain, spleen and lung that reflected expression levels of the human *PARK2* gene in the corresponding human organs. Indeed, the promoter of the transgene did consistently show tissue-specific activity in our experiments. The exact regulatory sequences in the promoter construct conferring this tissue-specificity in expression remained unknown. Nevertheless, the result showed that both mouse strains carried a transgene that was not ubiquitously expressed without regulation. At least, some of the sequences needed for tissue-specific expression of *PARK2* must be included within the promoter segment included in the construct. A likely location of some of the regulatory sequences is within the first 800 bp upstream of *PARK2*. This promoter segment is characterized by an H3K27Ac mark and the binding of over 30 transcription factors in the ENCODE cell lines.

While there is a high degree of sequence identity in the core promoter regions of both mouse and human *Parkin* genes, the identities of the transcriptional factors binding the mouse promoter are unknown. Hence, it is possible that the transcriptional factors controlling expression of the mouse endogenous *Park2* gene only partly overlap those that are required to trigger human *PARK2* activity. This could explain why in broad outline the reporter

construct displayed tissue-specificity in the mouse but only poor correlation with the expression levels of the endogenous gene.

Although very useful, one downfall of the *Hprt* transgenesis method is that the gene and its promoter are not being observed in their natural environment. Indeed, human *PARK2* is located on chromosome 6[80], as the homologous gene in the mouse is located on chromosome 17[110]. The X chromosome may not be a favorable environment to accurately survey promoter activity and even less the variants impacting this regulatory element. Other regulatory elements in cis or trans position (close or at a distance) of the endogenous gene may influence the power of the polymorphisms in the core promoter region[113].

Even though we did not aim to investigate a sex effect related to Parkin, we used both sexes in our studies to maximize data collected. At times, we observed significant differences in pPARK2/FLuc construct expression when comparing males and females. This sex effect was most pronounced with LPS challenge. Technically, female transgenic mice have two copies of the transgene. However, due to X-inactivation we did not expect any problem with gene dosage since females should have only one functional copy of the transgene[101]. Interestingly, the divergence between males and females was mainly observed to immune stimulation. As best shown with LPS challenge, only females repressed luciferase activity. It is well known for humans that males and females can display striking differences in susceptibility to infectious diseases. Likewise, susceptibility of the mouse to *Salmonella* infection also showed strong sex

effects[114]. However, the factors involved in the mediation of the sex specific repression of the human *PARK2* construct in the mouse are entirely unknown.

The systematic down-regulation of *Park2* after in vivo immune stimulation with various agents (bacteria and endotoxin) is a step toward a better understanding of the interaction between Parkin and host defenses. By taking into consideration the known function of Parkin as an E3 ubiquitin ligase, we can hypothesize that its down-regulation is necessary to promote the expression of needed cytokines and chemokines for the host clearance of the pathogen (Fig. 4.1). This pathway has been hypothesized by Kang et al., 2011 through the NFκB complex in mouse primary microglia and macrophages[5]. Possibly, at its constitutive levels, Parkin limits the production of these immune molecules to prevent chronic inflammation. Therefore, proper repression of the gene would be necessary to trigger an immune response in case of infection. On the other hand, de Léséleuc et al., have demonstrated that *Parkin* silencing reduces interleukin 6 (IL-6) and monocyte chemoattractant protein 1 (MCP-1) production in THP-1 macrophages, human monocyte-derived macrophages and human Schwann cells after stimulation with LPS or mycobacteria with a mechanism that is independent of the TLR-NFκB pathway[115]. Also, if *PARK2* repression is triggered bypathogen invasion, it could imply that Parkin is involved in tagging bacterial protein for pathogen destruction.

Regardless, there was a striking difference in *Park2* repression depending on the pathogen or endotoxin. Although the impact of BCG infection on *Park2* expression was modest, the effects with LPS challenge and *S. Typhimurium*

infection were large. Even though the BCG infected samples did not reach statistical significance in terms of *Park2* repression, we cannot disregard the potential biological significance of these results. It is possible that down-regulation of murine *Park2* is correlated to the pathogenicity of the parasite. Indeed, C56BL/6 mice are quite susceptible to *Salmonella* due to a mutation at the *Slc11A1* locus[116], but BCG Russia is generally considered a attenuated mycobacterial strain with low virulence[117]. As for LPS, even though it is not a pathogen, it is a very potent endotoxin from a gram-negative pathogen, *E. coli* which can cause septic choc in mice when injected directly into the blood stream[118]. It is also possible that the level of repression is influenced by the time point and that there is a time sensitivity regulating *Park2* repression after infection.

It is evident from the data collected in the course of this study that the *PARK2* gene is not only of interest for leprosy pathogenesis, but also for other infectious diseases. However, the specific role of Parkin in host defense remains to be described. At this point, we do not even know if Parkin down-regulation is a pathogen mechanism to overcome host defenses or if it is part of the host defense mechanism to clear bacteria and other pathogens from the body.

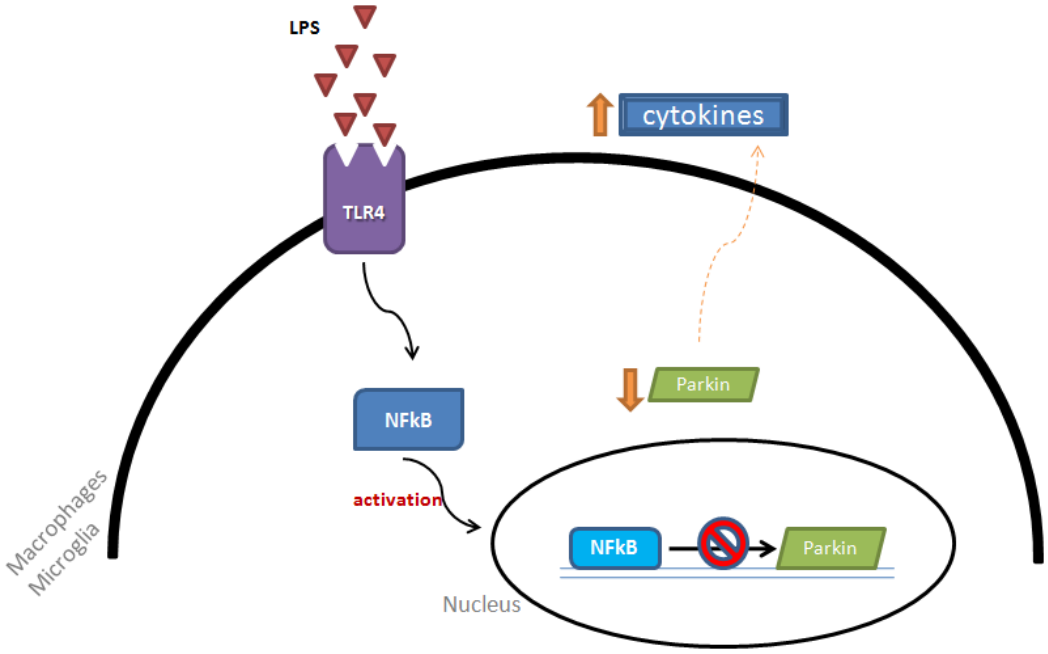

Figure 4.1 Possible pathway for Parkin involvement in host defense
The following model hypothesizes Parkin's interaction in microglia and macrophages. LPS binds to the TLR4 receptor which will lead to subsequent activation of the NFκB. NFκB than binds to the *Park2* promoter and may act as a repressor which down-regulates the gene expression. Through an unknown mechanism, it is possible that the repression of Parkin causes an augmentation of cytokines responsible for triggering other immune responses[5].

A main purpose of this work was to study the impact of two different human *PARK2* promoter allelic combinations on the gene expression. There are disorders where the abnormal sequence of a promoter is responsible for human disease. One known example is beta-thalassemia where the *HBB* gene, coding for beta-globin, an essential protein of hemoglobin, contains a polymorphism reducing its transcription. The regulatory SNPs causing disease can be located at several transcription factor binding sites[61]. Of course, the example provided previously is from a Mendelian disorder. However, genetically controlled differences in gene expression are assumed to be critically important in complex traits. It is intuitively understandable that metabolic pathways may be

considerably impacted by repressing or enhancing a key gene and that this may be reflected in changed susceptibility to complex diseases[113].

The results obtained in the transgenic mouse model did not support a functional role of the two tested polymorphisms. While differences in the gene regulation between mouse and human *Park2/PARK2* genes may be one reason, it is also possible that the tested polymorphisms are not causally related to leprosy susceptibility. Both SNPs are part of much larger SNP bins, i.e. groups of highly correlated markers. In fact, on statistical grounds it is not possible to give preference to any of the SNPs in the same bin to be the leprosy susceptibility factor. This implies that while the two SNPs tested represent the best multivariate model, other models are not significantly worse. Of interest for the selection of the promoter risk SNPs is a recent comparative SNP mapping study performed in Indian and Vietnamese leprosy patients[67]. Ultrahigh density SNP mapping in the two populations revealed that SNPs rs1040079 and rs2276201 (which was included in the construct) are not causally linked to leprosy risk. This was shown by the observation that the two SNPs segregated into an independent bin in the Indian population that was no longer associated with leprosy. In contrast, SNP rs9356058 was shown to be a leprosy risk factor in both the Indian and Vietnamese population supporting causality of this SNP for leprosy susceptibility. These data suggest that the construct used in our study may not have been optimally designed to find allele-specific effects on *PARK2* expression.

In theory, the only difference between the R3 and the S5 strains was the human *PARK2* promoter allelic combination. However, phenotypic differences

were noticed between mice of the R3 and S5 strains. A substantial difference in breeding efficiency was noticed in the course of our experiments since R3 mice reproduced more actively than S5 mice. As well, there was a difference in size of mice when comparing the two strains. S5 mice were smaller than their R3 counterparts. Finally, we noticed a dissimilarity in fur texture. Mice from the R3 strain had a rougher coat compared to mice of the S5 strain. No extensive studies were undertaken to identify possible additional disparities. The differences may be attributed to inbred variants accumulated over time in each strain. Such randomly fixed genetic differences could have made it difficult to evaluate the impact of the allelic combinations studied[103]. Ideally, it would have been preferable to generate and use more than one strain for each allelic combination to overcome such putative inbred variant effects. While this was the initial plan, such duplicate strains were lost during breeding.

Based on our results, it is tempting to conclude that the impact of *PARK2* SNPs on Parkin function is best studied in humans. For example, testing the impact of the allelic combinations of *PARK2/PACRG* leprosy susceptibility variants on *PARK2* levels derived dendritic cells stimulated with various bacteria including *M. leprae* might be an option. Initial results of experiments with whole blood assays showed that down-regulation of *PARK2* was detectable after bacterial stimulation (Schurr et al, unpublished data). Moreover, *PARK2* promoter variants were associated with the production of IL6 and CCL2 cytokines suggesting a possible link of those variants with leprosy susceptibility[115]. Yet, it is important to realize that all human studies by necessity are correlative in nature

and that it will not easily be possible to establish causality in humans. Hence, the identification of causal chains linking human genetic variation with leprosy susceptibility will still require animal experiments. Based on the results presented here, it is, however, critically important to have a very detailed understanding of the correlative matrices of human leprosy risk SNPs before initiating further mouse studies.

CONCLUSION

Overall, the pPARK2/FLuc constructs were expressed in mice at variable levels across organs indicating that the promoter region of the constructs was a functional and gene expression modulating regulatory region. The lack of correlation between expression of constructs and endogenous mouse *Park2* showed that the *Parkin* gene promoters did not behave identically in both species. We also determined that pPARK2/FLuc constructs and endogenous *Park2* expression were repressed in mouse spleens following infection with BCG and *S. Typhimurium.* Likewise, *Park2* was specifically down-regulated in the spleen when mice were challenged with LPS.

Differential expression between the two promoter constructs could not be validated partly due to large variability of reporter activity across experiments. Both mouse strains may have accumulated inbred variants that conceivably interfered with pPARK2/FLuc expression. We were unable to confirm our hypothesis that the two *PARK2* promoter constructs drive variable *PARK2* expression in the transgenic mouse model.

METHODS

Transgenic mice

S5 and R3 transgenic mice contain a transgene carrying alleles of the human *PARK2* promoter previously associated with leprosy susceptibility and resistance respectively[66] and a *FLuc* reporter gene coding for firefly luciferase. These mice were engineered from C56BL/6-129S embryonic stem cells by Alan Peterson using the *Hprt* locus transgenesis technique and were backcrossed to C57BL/6J mice to establish a fully inbred line. Breeding was maintained at the Montreal General Hospital and all experiments were performed following McGill University animal ethics policies.

BCG culture

Frozen stock of BCG Russia (American Type Culture Collection) was thawed and cultured in Middlebrook 7H9 medium (Difco Laboratories) supplemented with albumin-dextrose-catalase (Becton Dickinson and Co.) and 10% Tween 80 (Sigma-Aldrich). Cultures at an $OD_{600}$of 0.4 were diluted serially to obtain a culture of 10^6 colony forming units (CFU)/ml for injection. Injection doses were confirmed by plating different BCG dilutions on Middlebrook 7H10 agar (Difco Laboratories) supplemented with oleic acid-albumin-dextrose-catalase enrichment (Becton Dickinson and Co.) and BACTEC PANTA PLUS antibiotics (Becton Dickinson and Co.).

Salmonella Typhimurium culture

The *S. Typhimurium* strain Keller used in our experiments was kindly provided by Danielle Malo (McGill University, Montreal). Briefly, a small volume of bacterial suspension was plated on trypic soy agar (TSA) for colony isolation. The next day, one colony of *S. Typhimurium* was transferred into 5 ml of trypic soy broth (TSB, Becton and Dickson and Co.) and grown on a rotating platform at 37°C. The following day, 1ml of the inoculum was transferred to 100ml of TSB, incubated at 37°C until an OD_{600} of 0.1-0.2 was reached and then placed on ice for at least 2hours. The culture was then diluted to 10^5 CFU/ml for infection, and plated for CFU enumeration.

Bacterial infection and LPS challenge

S5 and R5 mice were challenged with BCG Russia, *S. Typhimurium*, or LPS. Mice were infected intravenously with a dose of approximately 10^5 CFU in 100 µl of PBS for BCG and 1000 CFU in 200ul of PBS for *S. Typhimurium*. LPS injections were performed through the intra-peritoneal route using 100 µg of LPS from *E. Coli* 055:B5 or *E. Coli* K-235 (Sigma-Aldrich) diluted in 0.5 ml of PBS. Five-weeks, 3-day and 48 hours were selected as the respective end points for BCG Russia, *S. Typhimurium*, and LPS challenge experiments. On the day of sacrification, infected mice and their controls were euthanatized by CO_2 asphyxiation and brains, lungs and spleens were collected. A fraction of the organs was placed RNA later (Qiagen) for subsequent RNA extraction or in lysis

buffer (PBS, 0.1% Triton-X-100 (Sigma-Aldrich) and cocktail of protease inhibitors (Roche)) to measure luciferase activity

Measurement of luciferase activity

Organs in Triton-X-100 lysis buffer were homogenized using an OMNI TH homogenizer. Clumps were removed from the homogenates first by sedimentation and then by filtration using Spin-X purification columns (Corning). To measure luciferase activity, 50 µl of reconstituted luciferase assay substrate from the Promega luciferase assay kit (Promega) was added to 10 µl purified lysate from each organ. The relative light units (RLU) for each organ were then measured with the Wallac luminometer. RLU values were normalized to total protein levels which were obtained using the Pierce Bicinchoninic acid (BCA) protein assay (Thermo Fisher Scientific). Briefly, purified homogenates were diluted by a factor of 20 of which 20 µl was placed on a 96-wellplate. 200 µl of working reagent was added to each well and the plate was incubated for 30 minutes at 37°C and protein concentration was measured with an ELISA microplate reader.

RNA extraction

The RNeasy Mini kit and Lipid Tissue kit (Qiagen) were used for RNA extraction from spleen and brain respectively following the manufacturer's protocols. The concentration and integrity of all RNA samples was verified using the NanoDrop ND-1000 spectrophotometer and the Agilent Bioanalyzer 2100.

Gene quantification studies.

Synthesis of first-strand cDNA was performed with 1μg of total RNA from mouse brain and spleen using the QuantiTect reverse transcription kit (Qiagen), as recommended by the manufacturer. For quantification studies, TaqMan Gene Expression assays for *Park2* (conjugated to the FAM fluorophore) and the *Gapdh* endogenous control gene (conjugated to the JOE fluorophore) were purchased from Applied Biosystems (Life Technologies). cDNA was amplified using Fermentas reagents on the Rotor-Gene 3000 (Corbett Research, Sydney, Australia), as specified.

Statistical analysis

The luciferase (RLU) and total protein (μg) ratio results were analyzed using unpaired Student's t-test with Prism Software to find statistical differences between groups. The correlation study between luciferase assay quantification over total protein and *Park2* expression was done using the RT-PCR Ct values associated to its respective RLU/μg of total protein value for each sample. For each pathogen shown, control mice were pooled with stimulated mice for correlation calculations. Fold changes in transcription levels between infected and control samples were determined by the threshold cycle (ΔΔCT) calculation[119]. Data are shown on a logarithmus dualis ($\log_2$) scale as the mean fold differences ± the standard error of the mean (SEM), which assumes an optimum PCR efficiency (E) of 2[120]. To perform fold difference calculations in instances where the *Park2* levels were completely undetectable following infection, missing values were

replaced by the highest detectable Ct value for the entire run. P values of < 0.05 were used to indicate statistical significance for all experiments.

BIBLIOGRAPHY

1. Hristova, V.A., Beasley, S.A., Rylett, R.J. & Shaw, G.S. Identification of a novel Zn2+-binding domain in the autosomal recessive juvenile Parkinson-related E3 ligase Parkin. *J Biol Chem* **284**, 14978-86 (2009).
2. Zhang, F.-R. *et al.* Genomewide Association Study of Leprosy. *N Eng J Med* **361**, 2609-2618 (2009).
3. Vissa, V.D. & Brennan, P.J. The genome of Mycobacterium leprae: a minimal mycobacterial gene set. *Genome Biol***2**, Review 1023 (2001).
4. Cole, S.T. *et al.* Massive gene decay in the leprosy bacillus. *Nature* **409**, 1007-11 (2001).
5. Tran, T.A. *et al.* Lipopolysaccharide and tumor necrosis factor regulate Parkin expression via nuclear factor-kappa B. *PLoS One***6**, e23660 (2011).
6. Rodrigues, L.C. & Lockwood, D.N.J. Leprosy now: epidemiology, progress, challenges, and research gaps. *The Lancet Infectious Diseases* **11**, 464-470 (2011).
7. Walden, H. & Martinez-Torres, R.J. Regulation of Parkin E3 ubiquitin ligase activity. *Cell Mol Life Sci* **69**, 3053-67 (2012).
8. Schurr, E., Alcais, A., de Leseleuc, L. & Abel, L. Genetic predisposition to leprosy: A major gene reveals novel pathways of immunity to Mycobacterium leprae. *Semin Immunol* **18**, 404-10 (2006).
9. Misch, E.A., Berrington, W.R., Vary, J.C., Jr. & Hawn, T.R. Leprosy and the human genome. *Microbiol Mol Biol Rev* **74**, 589-620 (2010).
10. Abulafia, J. & Vignale, R.A. Leprosy: pathogenesis updated. *Int J Dermatol***38**, 321-34 (1999).
11. Hansen, G.H.A. Investigations concerning the etiology of leprosy. *Norsk Magazin for Lægevidenskaben* **4**, 88 (1874).
12. Matheson, C.D. *et al.* Molecular exploration of the first-century Tomb of the Shroud in Akeldama, Jerusalem. *PLoS One* **4**, e8319 (2009).
13. Hulse, E.V. Leprosy and ancient Egypt. *Lancet* **2**, 1024-5 (1972).
14. Liveing, R. *Elephantiasis Graecorum: Or True Leprosy (1873)*, (Kessinger Publishing, 2008).
15. Navon, L. Beggars, metaphors, and stigma: a missing link in the social history of leprosy. *Soc Hist Med***11**, 89-105 (1998).
16. Haas, L.F. Armauer Gerhard Heinrik Hansen (1841-1912). *J Neurol Neurosurg Psychiatry* **67**, 14 (1999).
17. Cottle, W. Chaulmoogra Oil in Leprosy. *Br Med J* **1**, 968-9 (1879).
18. Jacobson, R.R. Treatment of leprosy. in *Leprosy* (ed. R.H., I.H.) 32 (Churchill Livingston, Edimburg, 1995).
19. Scollard, D.M. Treatment gets better, but leprosy remains a global problem. *Ann Acad Med Singapore* **37**, 1-2 (2008).
20. Shepard, C.C. The experimental disease that follows the injection of human leprosy bacilli into foot-pads of mice. *J Exp Med* **112**, 445-54 (1960).
21. Rees, R.J., Pearson, J.M. & Waters, M.F. Experimental and clinical studies on rifampicin in treatment of leprosy. *Br Med J***1**, 89-92 (1970).
22. Jacobson, R.R. & Hastings, R.C. Rifampin-resistant leprosy. *Lancet* **2**, 1304-5 (1976).

23. Matsuoka, M. Drug resistance in leprosy. *Jpn J Infect Dis* **63**, 1-7 (2010).
24. Cholo, M.C., Steel, H.C., Fourie, P.B., Germishuizen, W.A. & Anderson, R. Clofazimine: current status and future prospects. *J Antimicrob Chemother* **67**, 290-8 (2012).
25. Chemotherapy of leprosy for control programmes. *World Health Organ Tech Rep Ser* **675**, 1-33 (1982).
26. WHO donated MDT. in *Leprosy elimination* (World health organisation).
27. Chemotherapy of leprosy. Report of a WHO Study Group. *World Health Organ Tech Rep Ser* **847**, 1-24 (1994).
28. Scollard, D.M. *et al.*The continuing challenges of leprosy. *Clin Microbiol Rev***19**, 338-81 (2006).
29. Suzuki, K. *et al.*Current status of leprosy: epidemiology, basic science and clinical perspectives. *J Dermatol* **39**, 121-9 (2012).
30. Kaur, S., Sharma, V.K. & Kumar, B. Treatment of leprosy--newer concepts. *Indian J Lepr* **56**, 307-12 (1984).
31. Ridley, D.S. & Jopling, W.H. Classification of leprosy according to immunity. A five-group system. *Int J Lepr Other Mycobact Dis* **34**, 255-73 (1966).
32. Ridley, D.S. Histological classification and the immunological spectrum of leprosy. *Bull World Health Organ* **51**, 451-65 (1974).
33. Rambukkana, A., Salzer, J.L., Yurchenco, P.D. & Tuomanen, E.I. Neural targeting of Mycobacterium leprae mediated by the G domain of the laminin-alpha2 chain. *Cell* **88**, 811-21 (1997).
34. Rea, T.H. St. John Dermatological Society Oration, 1994. Immune responses in leprosy, cytokines and new archetypes for dermatology. *Clin Exp Dermatol***20**, 89-97 (1995).
35. World Health Assembly resolution 1991. (World Health Assembly).
36. WHO weekly record leprosy 2012. (World Health Organization)
37. Worobec, S.M. Treatment of leprosy/Hansen's disease in the early 21st century. *Dermatol Ther* **22**, 518-37 (2009).
38. Truman, R.W. *et al.* Probable zoonotic leprosy in the southern United States. *N Engl J Med* **364**, 1626-33 (2011).
39. Kerr-Pontes, L.R. *et al.*Socioeconomic, environmental, and behavioural risk factors for leprosy in North-east Brazil: results of a case-control study. *Int J Epidemiol* **35**, 994-1000 (2006).
40. Brosch, R., Gordon, S.V., Eiglmeier, K., Garnier, T. & Cole, S.T. Comparative genomics of the leprosy and tubercle bacilli. *Research in Microbiology* **151**, 135-142 (2000).
41. Pinheiro, R.O., de Souza Salles, J., Sarno, E.N. & Sampaio, E.P. Mycobacterium leprae-host-cell interactions and genetic determinants in leprosy: an overview. *Future Microbiol* **6**, 217-30 (2011).
42. Shepard, C.C. The first decade in experimental leprosy. *Bull World Health Organ* **44**, 821-7 (1971).
43. Microbiology of M. leprae. in *Leprosy elimination* Vol. 2012 (WHO, 2012).
44. Meyers, W.M., Gormus, B.J., Walsh, G.P., Baskin, G.B. & Hubbard, G.B. Naturally acquired and experimental leprosy in nonhuman primates. *Am J Trop Med Hyg* **44**, 24-7 (1991).

45. Hall, B.G. & Salipante, S.J. Molecular epidemiology of Mycobacterium leprae as determined by structure-neighbor clustering. *J Clin Microbiol* **48**, 1997-2008 (2010).
46. Barker, L.P. Mycobacterium leprae interactions with the host cell: recent advances. *Indian J Med Res* **123**, 748-59 (2006).
47. Brosch, R., Pym, A.S., Gordon, S.V. & Cole, S.T. The evolution of mycobacterial pathogenicity: clues from comparative genomics. *Trends in Microbiology* **9**, 452-8 (2001).
48. Walker, L., LeVine, H. & Jucker, M. Koch's postulates and infectious proteins. *Acta Neuropathologica* **112**, 1-4 (2006).
49. Clementi, M. & Di Gianantonio, E. Genetic susceptibility to infectious diseases. *Reprod Toxicol* **21**, 345-9 (2006).
50. Rowell, J.L. *et al.* Trends in Population-Based Studies of Human Genetics in Infectious Diseases. *PLoS ONE* **7**, e25431 (2012).
51. Borecki, I.B. & Province, M.A. Genetic and Genomic Discovery Using Family Studies. *Circulation* **118**, 1057-1063 (2008).
52. Jepson, A. Twin studies for the analysis of heritability of infectious diseases. *Bulletin de l'Institut Pasteur* **96**, 71-81 (1998).
53. Evangelou, E., Trikalinos, T.A., Salanti, G. & Ioannidis, J.P.A. Family-Based versus Unrelated Case-Control Designs for Genetic Associations. *PLoS Genet* **2**, e123 (2006).
54. Hill, A.V. Aspects of genetic susceptibility to human infectious diseases. *Annu Rev Genet* **40**, 469-86 (2006).
55. Frodsham, A.J. & Hill, A.V. Genetics of infectious diseases. *Hum Mol Genet* **13**, R187-94 (2004).
56. Newport, M.J. & Finan, C. Genome-wide association studies and susceptibility to infectious diseases. *Brief Funct Genomics* **10**, 98-107 (2011).
57. Fellay, J. *et al.* A whole-genome association study of major determinants for host control of HIV-1. *Science* **317**, 944-7 (2007).
58. Vannberg, F.O., Chapman, S.J. & Hill, A.V. Human genetic susceptibility to intracellular pathogens. *Immunol Rev* **240**, 105-16 (2011).
59. Hill, A.V. The immunogenetics of human infectious diseases. *Annu Rev Immunol* **16**, 593-617 (1998).
60. Chapman, S.J. & Hill, A.V. Human genetic susceptibility to infectious disease. *Nat Rev Genet* **13**, 175-88 (2012).
61. Epstein, D.J. Cis-regulatory mutations in human disease. *Brief Funct Genomic Proteomic* **8**, 310-6 (2009).
62. Amos, W., Driscoll, E. & Hoffman, J.I. Candidate genes versus genome-wide associations: which are better for detecting genetic susceptibility to infectious disease? *Proc Biol Sci* **278**, 1183-8 (2011).
63. Alter, A., Alcais, A., Abel, L. & Schurr, E. Leprosy as a genetic model for susceptibility to common infectious diseases. *Hum Genet* **123**, 227-35 (2008).
64. Mira, M.T. Genetic host resistance and susceptibility to leprosy. *Microbes Infect* **8**, 1124-31 (2006).
65. Orlova, M., Di Pietrantonio, T. & Schurr, E. Genetics of infectious diseases: hidden etiologies and common pathways. *Clin Chem Lab Med* **49**, 1427-37 (2011).

66. Mira, M.T. *et al.* Susceptibility to leprosy is associated with PARK2 and PACRG. *Nature* **427**, 636-40 (2004).
67. Alter, A. *et al.* Linkage disequilibrium pattern and age-at-diagnosis are critical for replicating genetic associations across ethnic groups in leprosy. *Hum Genet* (2012).
68. Malik, S. *et al.* Alleles of the NRAMP1 gene are risk factors for pediatric tuberculosis disease. *Proc Natl Acad Sci U S A* **102**, 12183-8 (2005).
69. Abel, L. *et al.* Susceptibility to leprosy is linked to the human NRAMP1 gene. *J Infect Dis* **177**, 133-45 (1998).
70. Ali, S. *et al.* PARK2/PACRG polymorphisms and susceptibility to typhoid and paratyphoid fever. *Clin Exp Immunol* **144**, 425-31 (2006).
71. Schurr, E. & Gros, P. A common genetic fingerprint in leprosy and Crohn's disease? *N Engl J Med* **361**, 2666-8 (2009).
72. Hakimi, M. *et al.* Parkinson's disease-linked LRRK2 is expressed in circulating and tissue immune cells and upregulated following recognition of microbial structures. *J Neural Transm* **118**, 795-808 (2011).
73. Fritsch, T. *et al.* Parkinson disease: research update and clinical management. *South Med J* **105**, 650-6 (2012).
74. Giasson, B.I. & Lee, V.M. Parkin and the molecular pathways of Parkinson's disease. *Neuron* **31**, 885-8 (2001).
75. Forno, L.S. Neuropathology of Parkinson's disease. *J Neuropathol Exp Neurol* **55**, 259-72 (1996).
76. Mizuno, Y., Hattori, N., Mori, H., Suzuki, T. & Tanaka, K. Parkin and Parkinson's disease. *Curr Opin Neurol* **14**, 477-82 (2001).
77. Humbert, J. *et al.* Parkin and synphilin-1 isoform expression changes in Lewy body diseases. *Neurobiol Dis* **26**, 681-7 (2007).
78. Kim, Y. *et al.* PINK1 controls mitochondrial localization of Parkin through direct phosphorylation. *Biochem Biophys Res Commun* **377**, 975-80 (2008).
79. Saiki, S., Sato, S. & Hattori, N. Molecular pathogenesis of Parkinson's disease: update. *J Neurol Neurosurg Psychiatry* **83**, 430-6 (2012).
80. Kitada, T. *et al.* Mutations in the Parkin gene cause autosomal recessive juvenile Parkinsonism. *Nature* **392**, 605-8 (1998).
81. Bonifati, V. Autosomal recessive Parkinsonism. *Parkinsonism Relat Disord* **18 Suppl 1**, S4-6 (2012).
82. Mori, H. *et al.* Pathologic and biochemical studies of juvenile Parkinsonism linked to chromosome 6q. *Neurology* **51**, 890-2 (1998).
83. Eisenhaber, B., Chumak, N., Eisenhaber, F. & Hauser, M.-T. The ring between ring fingers (RBR) protein family. *Genome Biology* **8**, 209 (2007).
84. Feany, M.B. & Pallanck, L.J. Parkin: a multipurpose neuroprotective agent? *Neuron* **38**, 13-6 (2003).
85. Kao, S.Y. Regulation of DNA repair by Parkin. *Biochem Biophys Res Commun* **382**, 321-5 (2009).
86. Chung, K.K. *et al.* S-nitrosylation of Parkin regulates ubiquitination and compromises Parkin's protective function. *Science* **304**, 1328-31 (2004).
87. Shiba-Fukushima, K. *et al.* PINK1-mediated phosphorylation of the Parkin ubiquitin-like domain primes mitochondrial translocation of Parkin and regulates mitophagy. *Sci Rep* **2**, 1002 (2012).

88. Chen, D. *et al.* Parkin mono-ubiquitinates Bcl-2 and regulates autophagy. *J Biol Chem* **285**, 38214-23 (2010).
89. Kuroda, Y., Mitsui, T., Kunishige, M. & Matsumoto, T. Parkin affects mitochondrial function and apoptosis in neuronal and myogenic cells. *Biochemical and Biophysical Research Communications* **348**, 787-793 (2006).
90. Cesari, R. *et al.* Parkin, a gene implicated in autosomal recessive juvenile Parkinsonism, is a candidate tumor suppressor gene on chromosome 6q25-q27. *Proc Natl Acad Sci U S A* **100**, 5956-61 (2003).
91. Mehdi, S.J., Ali, A. & Rizvi, M.M. Parkin gene alterations in ovarian carcinoma from northern Indian population. *Pathol Oncol Res* **17**, 579-86 (2011).
92. Fujiwara, M. *et al.* Parkin as a tumor suppressor gene for hepatocellular carcinoma. *Oncogene* **27**, 6002-11 (2008).
93. Kao, S.Y. DNA damage induces nuclear translocation of Parkin. *J Biomed Sci* **16**, 67 (2009).
94. Rothfuss, O. *et al.* Parkin protects mitochondrial genome integrity and supports mitochondrial DNA repair. *Hum Mol Genet* **18**, 3832-50 (2009).
95. LaVoie, M.J., Cortese, G.P., Ostaszewski, B.L. & Schlossmacher, M.G. The effects of oxidative stress on Parkin and other E3 ligases. *J Neurochem* **103**, 2354-68 (2007).
96. West, A.B., Lockhart, P.J., O'Farell, C. & Farrer, M.J. Identification of a novel gene linked to Parkin via a bi-directional promoter. *J Mol Biol* **326**, 11-9 (2003).
97. Imai, Y. *et al.* A product of the human gene adjacent to Parkin is a component of Lewy bodies and suppresses Pael receptor-induced cell death. *J Biol Chem* **278**, 51901-10 (2003).
98. Waterston, R.H. *et al.* Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520-62 (2002).
99. Di Pietrantonio, T. & Schurr, E. Mouse models for the genetic study of tuberculosis susceptibility. *Brief Funct Genomic Proteomic* **4**, 277-92 (2005).
100. Gharib, W.H. & Robinson-Rechavi, M. When orthologs diverge between human and mouse. *Brief Bioinform* **12**, 436-41 (2011).
101. Bronson, S.K. *et al.* Single-copy transgenic mice with chosen-site integration. *Proc Natl Acad Sci U S A* **93**, 9067-72 (1996).
102. Keeler, C.E. *The laboratory mouse; its origin, heredity, and culture, by Clyde E. Keeler*, (Harvard university press, Cambridge, (1931).
103. Keane, T.M. *et al.*Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* **477**, 289-94 (2011).
104. Di Pietrantonio, T., Correa, J.A., Orlova, M., Behr, M.A. & Schurr, E. Joint effects of host genetic background and mycobacterial pathogen on susceptibility to infection. *Infect Immun* **79**, 2372-8 (2011).
105. Schofield, P.N., Hoehndorf, R. & Gkoutos, G.V. Mouse genetic and phenotypic resources for human genetics. *Hum Mutat* **33**, 826-36 (2012).
106. Austin, C.P. *et al.* The knockout mouse project. *Nat Genet***36**, 921-4 (2004).
107. Vidal, S.M., Malo, D., Vogan, K., Skamene, E. & Gros, P. Natural resistance to infection with intracellular parasites: isolation of a candidate for Bcg. *Cell* **73**, 469-85 (1993).
108. Malo, D. *et al.* Haplotype mapping and sequence analysis of the mouse Nramp gene predict susceptibility to infection with intracellular parasites. *Genomics* **23**, 51-61 (1994).

109. Dietrich, W.F. Using mouse genetics to understand infectious disease pathogenesis. *Genome Res* **11**, 325-31 (2001).
110. Tomac, A.C. & Hoffer, B.J. Assignment of the mouse Park2 (PARKIN), the homologue to a new human Parkinson candidate gene, to the telomeric region of mouse 17A3.2-3.3, by in situ hybridization. *Cytogenet Cell Genet* **95**, 120-1 (2001).
111. Ramsey, C.P. & Giasson, B.I. Identification and characterization of a novel endogenous murine Parkin mutation. *J Neurochem* **113**, 402-17 (2010).
112. Romano, M. & Huygen, K. An update on vaccines for tuberculosis - there is more to it than just waning of BCG efficacy with time. *Expert Opin Biol Ther* **12**, 1601-10 (2012).
113. Maston, G.A., Evans, S.K. & Green, M.R. Transcriptional regulatory elements in the human genome. *Annu Rev Genomics Hum Genet* **7**, 29-59 (2006).
114. Caron, J. *et al.* Identification of genetic loci controlling bacterial clearance in experimental Salmonella enteritidis infection: an unexpected role of Nramp1 (Slc11a1) in the persistence of infection in mice. *Genes Immun* **3**, 196-204 (2002).
115. de Leseleuc, L. *et al.* PARK2 mediates interleukin 6 and monocyte chemoattractant protein 1 production by human macrophages. *PLoS Negl Trop Dis* **7**, e2015 (2013).
116. Swearingen, M.C., Porwollik, S., Desai, P.T., McClelland, M. & Ahmer, B.M.M. Virulence of 32 *Salmonella* Strains in Mice. *PLoS ONE* **7**, e36043 (2012).
117. Leversen, N.A., Sviland, L., Wiker, H.G. & Mustafa, T. Long-Term Persistence of BCG Pasteur in Lungs of C57BL/6 Mice Following Intranasal Infection. *Scandinavian Journal of Immunology* **75**, 489-499 (2012).
118. Khan, N.A. *et al.* Mitigation of septic shock in mice and rhesus monkeys by human chorionic gonadotrophin-related oligopeptides. *Clin Exp Immunol* **160**, 466-78 (2010).
119. Livak, K.J. & Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-8 (2001).
120. Fu, W.J., Hu, J., Spencer, T., Carroll, R. & Wu, G. Statistical models in assessing fold change of gene expression in real-time RT-PCR experiments. *Comput Biol Chem* **30**, 21-6 (2006).