Exploring host and pathogen biomarkers for leprosy

Tio Coma, M.

English summary
Nederlandse samenvatting
Acknowledgements
Curriculum vitae
List of publications

English summary

Leprosy is a multifactorial chronic disease caused by *Mycobacterium leprae* or *Mycobacterium lepromatosis* that affects the skin and nerves. More than 200.000 new cases are diagnosed per year; thus, transmission is still ongoing. The most likely way of transmission is the respiratory route form human-to-human; however, transmission is still not clearly understood. Early diagnosis of leprosy is crucial to reduce and avoid transmission as well as leprosy-associated disabilities, which are also a cause of stigma. Currently, diagnosis is performed based on clinical signs and symptoms and late- or mis-diagnosis are not uncommon.

In this thesis, we combined the study of pathogen transmission with host transcriptomic and genomic biomarkers. To explore *M. leprae* transmission a One Health approach was followed, where human, animal and environmental samples were studied.

In **chapter 2**, *M. leprae* transmission in multibacillary leprosy patients as well as their household contacts (HC) was studied in Bangladesh. We observed that *M. leprae* was not only present in leprosy patients, but also in asymptomatic individuals. Particularly in the nasal cavities, with up to 18% of asymptomatic HC showing presence of *M. leprae* DNA and 36.8% with phenolic glycolipid I (PGL-I) IgM in serum.

M. leprae whole genomes have been retrieved over the years and a genotype classification was established consisting of four genotypes (1-4) and 16 subtypes (A-P). In **chapter 2**, *M. leprae* genome diversity in Bangladesh was explored and several subtypes of genotype 1 (1A, 1C and 1D) were identified with subtype 1D being the most prevalent. Importantly, we describe a new subtype, 1B-Bangladesh, only found in Bangladesh up until now. Moreover, we demonstrated that the subtype 1C does not constitute a separate subtype and is part of genotype 1D.

In **chapter 3**, we investigated whether soil could be a potential reservoir of the leprosy bacilli. We identified *M. leprae* DNA in soil from the houses of leprosy patients in Bangladesh and from armadillos' holes at former leprosy colonies in Suriname. Additionally, *M. lepromatosis* was detected in soil obtained around an area where *M. lepromatosis* infected squirrels are located in the Isle of Arran. Nevertheless, the low concentration of *M. leprae* or *M. lepromatosis* DNA found in soil suggests that environmental contamination as a source of infection is not very likely.

Leprosy is not exclusive to humans, as red squirrels, armadillos and non-human primates can become naturally infected with *M. leprae* or *M. lepromatosis*. Squirrels infected with *M. leprae* or *M. lepromatosis* were identified in the British Isles. In **chapter 4,** we investigated whether Dutch and Belgian squirrels were carriers of the leprosy bacilli. We examined 114 squirrels by quantitative PCR (qPCR) and we did not detect *M. leprae* or *M. lepromatosis* DNA in any of the animals. This is in line with previous findings in France, Germany, Swit-

zerland and Italy where PCR showed no presence of the leprosy bacilli either. Thus, up to the present time, squirrels infected with *M. leprae* or *M. lepromatosis* have only been identified in the British Isles. Transmission of the leprosy bacilli from red squirrels to humans has not been reported and due to limited squirrel-human interaction in the areas where squirrels with leprosy were found and the geographical limitations of islands, zoonotic transmission and transmission to other areas should not be a cause of major concerns. Nevertheless, it remains necessary to keep vigilant and include the study of animal and environmental reservoirs in strategies to effectively stop transmission.

Next, we searched for transcriptomic host biomarkers that could predict leprosy or leprosy reactions before occurrence of symptoms. In **chapter 5**, we aimed to develop a predictive transcriptomic signature in blood that could assess whether an individual intensely exposed to *M. leprae* would develop leprosy. Gene expression differences between leprosy progressors and HC who remained without symptoms were studied using RNA-Seq. Minimal longitudinal intra-individual variation was found in gene expression of leprosy progressors between the pre-symptomatic phase and the time of clinical diagnosis of leprosy. This indicates that gene expression differences between healthy individuals and those who will develop leprosy can be observed months before clinical symptoms are visible. A 4-gene transcriptomic signature, designated RISK4LEP, was identified and validated in Bangladesh. This RNA signature could identify HC of leprosy patients who developed leprosy before symptoms were visible, 4 to 61 months before clinical diagnosis. A machine learning algorithm, random forest, was used to identify the signature which was validated by reverse transcription quantitative PCR (RT-qPCR). RISK4LEP is based on the expression of 4 genes: *MT-ND2*, *REX1BD*, *TPGS1* and *UBC*.

Leprosy reactions are episodes of increased inflammation occurring unpredictably before, during or after multidrug treatment (MDT). Reactions are often late- or misdiagnosed, which may result into permanent neuropathy or disabilities caused by ulcers and other recurrent pathologies. In **chapter 6**, we employed dual color Reverse-Transcription Multiplex Ligation-dependent Probe Amplification (dcRT-MLPA) to identify a 5-gene signature which predicted leprosy reversal reactions (RR) ≥ 2 months before onset in leprosy patients from four endemic areas: Bangladesh, Brazil, Ethiopia and Nepal. The 5-gene RR signature is formed by *CCL2, CD8A, IL2, IL15* and *MARCO*. Applying our 5-gene RR signature in a POC test, would allow reduction of reaction-associated neuropathy.

Finally, in **chapter 7**, we investigated genomic biomarkers to identify individuals at higher risk of developing leprosy. Three genetic markers (rs1801224, rs13001714 and rs1801582) were associated with leprosy in a family study analysis in Bangladesh. These markers were previously described to be associated with leprosy in the Prata Village. Although the results were not replicated in a separate control-case set from the same area in Bangladesh,

our results are a validation of the association of these SNPs with leprosy in a distinct population. The described genetic markers are located at *CUBN*, *IL1RL1* and *PRKN* (formerly known as *PARK2*) genes. Genetic variants associated with susceptibility to leprosy can be incorporated into a genetic profile to identify individuals exposed to *M. leprae* who present a higher risk of developing leprosy.

The combination of demographic characteristics, pathogen detection, genetic and/or transcriptomic biomarkers can be applied in a multifactorial leprosy signature applicable for early diagnosis of leprosy and/or to guide intervention strategies. Identification of predictive biomarkers will in due course lead to prompt treatment, preventing leprosy-associated irreversible disabilities as well as reducing *M. leprae* transmission.