Analysis of the Role of Host Genetics in Shaping Diversity of the Murine Lung Microbiota

Dissertation
in fulfilment of the requirements for the degree
Doctor rerum naturalium
of the Faculty of Mathematics and Natural Sciences
at the Christian Albrechts University of Kiel

Submitted by Juryung Chung

Department of Evolutionary Genetics,
Max Planck Institute for Evolutionary Biology

Plön, 2021

First referee: Prof. Dr. John F. Baines

Second referee: Prof. Dr. Hinrich Schulenburg

Additional referee: Prof. Dr. Thomas Roeder

Date of oral examination: 28.06.2021

Signature: _____

Table of Contents

Affidavit

Written declaration/Erklärung

Hereby I declare that

i. apart from my supervisor's guidance, the content and design of this dissertation is the product of my own work and only using the sources listed. The co-author's contributions to specific paragraphs are listed in the thesis outline section;
ii. this thesis has not already been submitted either partially or wholly as part of a doctoral degree to another examination body, and no other materials are published or submitted for publication than indicated in the thesis;
iii. the preparation of the thesis has been subjected to the Rules of Good Scientific Practice of the German Research Foundation;
iv. an academic degree has never been withdrawn.

Author contributions

Chapter 1: Prof. Dr. John Baines and Prof. Dr. Saleh Ibrahim designed the study. Dr. Meriem Belheouane dissected the mice. Dr. Sven Künzel performed the MiSeq sequencing. Shauni Doms designed the pipeline for processing the sequences and Dr. Yask Gupta and Dr. Meriem Belheouane designed the pipeline for QTL mapping. I performed all laboratory experiments and data analyses, wrote the chapter with editing from Prof. Dr. John Baines.

Chapter 2: Dr. Danielle Harris provided a *Pelomonas* sample along with valuable discussions on the bacteria. I performed all laboratory experiments and data analyses, and wrote the chapter with editing from Prof. Dr. John Baines.

Chapter 3: Prof. Dr. John Baines and I designed the study. I planned the overall mouse breeding scheme with help from Dr. Sven Künzel. Dr. Marie Vallier and I dissected mice and preserved the tissues. Dr. Sven Künzel performed the MiSeq sequencing. I performed all

laboratory experiments, including genotyping and data analyses, and wrote the chapter with editing from Prof. Dr. John Baines.

Chapter 4: Prof. Dr. John Baines designed the study. Dr. Marie Vallier, Aleksa Cèpic, and I dissected the mice and preserved the tissues. Dr. Sven Künzel performed the MiSeq sequencing. I performed all laboratory experiments and data analyses, and wrote the chapter with editing from Prof. Dr. John Baines.

Plön, April 2021,

Juryung Chung	Prof. Dr. John F. Baines

Abstract

Although lungs were long considered as sterile, the presence and importance of microbes inhabiting the lungs are now widely recognized. Using the 15th generation ("G_{15}") of a mouse advanced intercross line (AIL) population, we performed QTL mapping in an effort to identify host genes that influence lung microbes and may play a role in lung functioning and disease susceptibility. The lung microbiota of the G_{15} mouse population was examined through bacterial 16S rRNA gene amplicon sequencing, whereby *Lactobacillus* was identified as the most abundant genus. High-precision droplet digital PCR (ddPCR) was adapted in order to refine the dataset, remove contaminants, and quantify the absolute load of the candidate genera *Lactobacillus* and *Pelomonas*. The association between host genetic loci and bacterial traits was determined using a QTL linkage mapping approach on the relative abundances of the core microbiota as well as the absolute load data of the two candidate genera. This analysis yielded candidate host genes associated with various lung diseases. For example, two candidate genes, *Il-10* and *Mk2*, were detected from mapping *Lactobacillus* load and were especially interesting as both are involved in regulating inflammatory responses. As such, these genes are known to be involved in lung diseases such as lung cancer, cystic fibrosis, and allergic airway inflammation. To test for potential interaction and feedback between these genes and *Lactobacillus*, their gene expression levels were measured using ddPCR, which revealed a significant negative correlation between *Lactobacillus* load and *Mk2* expression. A follow-up study was performed with *Il-10* knockout (KO) mice, but the initial results revealed no significant differences in lung microbiota composition according to *Il-10* genotype. Finally, a candidate gene analysis was performed in wild type versus *B4galnt2* knock out mice, as this blood group related gene was previously shown to influence the gut microbiota, and is also predicted to be expressed in the lung. This analysis also failed to detect an influence of host genotype. Through this in-depth study of the murine lung microbiota, we generated novel working hypotheses for the impact of candidate bacteria and disease genes on host inflammatory responses in the lung, which will be explored in future work.

Zusammenfassung

Obwohl die Lunge lange Zeit als steril angesehen wurde, sind das Vorhandensein und die Bedeutung von Mikroben, die die Lunge besiedeln, heute allgemein anerkannt. Unter Verwendung der 15. Generation ("G15") einer Mauspopulation einer fortgeschrittenen Kreuzungslinie (AIL) führten wir eine QTL-Kartierung durch, um Wirtsgene zu identifizieren, die die Lungenmikroben beeinflussen und möglicherweise eine Rolle bei der Lungenfunktion und Krankheitsanfälligkeit spielen. Die Lungenmikrobiota der G15-Mauspopulation wurde mittels bakterieller 16S rRNA-Gen-Amplikon-Sequenzierung untersucht, wobei *Lactobacillus* als die an der häufigsten vorkommenden Gattung identifiziert wurde. Hochpräzise digitale Tröpfchen-PCR (ddPCR) wurde angepasst, um den Datensatz zu verfeinern, Verunreinigungen zu entfernen und die absoluten Mengen der Kandidatengattungen *Lactobacillus* und *Pelomonas* zu quantifizieren. Die Assoziation zwischen genetischen Wirtsloci und bakteriellen Merkmalen wurde mithilfe eines QTL-Linkage-Mapping-Ansatzes auf den relativen Häufigkeiten der Kernmikrobiota sowie den absoluten Mengen-Daten der beiden Kandidatengattungen bestimmt. Diese Analyse ergab Kandidaten-Wirtsgene, die mit verschiedenen Lungenkrankheiten assoziiert sind. Zum Beispiel wurden zwei Kandidatengene, *Il-10* und *Mk2*, bei der Kartierung der *Lactobacillus*-Menge entdeckt und waren besonders interessant, da beide an der Regulierung von Entzündungsreaktionen beteiligt sind. Von diesen Genen ist bekannt, dass sie an Lungenerkrankungen wie Lungenkrebs, zystischer Fibrose und allergischer Atemwegsentzündung beteiligt sind. Um eine mögliche Interaktion und Rückkopplung zwischen diesen Genen und *Lactobacillus* zu testen, wurden ihre Genexpressionslevel mittels ddPCR gemessen, was eine signifikante negative Korrelation zwischen der *Lactobacillus*-Menge und der Mk2-Expression ergab. Eine Folgestudie wurde mit *Il-10*-Knockout (KO)-Mäusen durchgeführt, aber die ersten Ergebnisse zeigten keine signifikanten Unterschiede in der Zusammensetzung der Lungenmikrobiota bezüglich des *Il-10*-Genotyps. Schließlich wurde eine Kandidatengen-Analyse in Wildtyp- und *B4galnt2*-KO-Mäusen durchgeführt, da dieses blutgruppenverwandte Gen zuvor gezeigt wurde, dass es die Darm-Mikrobiota beeinflusst und auch in der Lunge exprimiert werden sollte. Auch bei dieser Analyse konnte kein Einfluss des Wirtsgenotyps festgestellt werden. Durch diese eingehende Studie des Lungenmikrobioms von Mäusen haben wir neue Hypothesen für den Einfluss von

Bakterienkandidaten und Krankheitsgenen auf die Entzündungsreaktionen des Wirts in der Lunge generiert, die in zukünftigen Arbeiten untersucht werden sollen.