Characterization of Skin-Resident Microbiota in Inflammatory Cutaneous Disease

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

Submitted by

Britt Marie Hermes
from California, USA

Kiel, 2023

First reviewer: Prof. Dr. John F. Baines
Second reviewer: Prof. Dr. Hinrich Schulenburg
Date of oral examination: October 6, 2022

Nothing in biology makes sense, except in the light of evolution.
-Theodosius Dobzhansky, 1973

Table of Contents

Declaration

I hereby declare that,

i. apart from my supervisor's guidance, the content and design of this thesis is completely my own work. Contributions of other authors are listed under author contributions.
ii. this thesis has not been submitted, either partially or completely, as part of a doctoral degree to another examining institution. No materials are published or submitted for publication other than indicated in this thesis.
iii. this thesis was prepared in compliance with the "Rules of Good Scientific Practice" of the German Research Foundation (DFG).
iv. I have not had an academic degree revoked.

Britt Marie Hermes

Britt Marie Hermes
Kiel, July 12, 2022

Prof. Dr. John F. Baines

Author contribution statements

Chapter 1
Belheouane, M.*, **Hermes, B. M.*,** Van Beek, N., Benoit, S., Bernard, P., Drenovska, K., Gerdes, S., Gläser, R., Goebeler, M., Günther, C., von Georg, A., Hammers, C. M., Holtsche, M. M., Homey, B., Horváth, O. N., Hübner, F., Linnemann, B., Joly, P., Márton, D., Patsatsi A., Pföhler C., Sárdy M., Huilaja L., Vassileva S., Zillikens D., Ibrahim S., Sadik C. D., Schmidt E., Baines, J. F. Characterization of the skin microbiota in bullous pemphigoid patients and controls reveals novel microbial indicators of disease. ***Journal of Advanced Research***. (2022). https://doi.org/10.1016/j.jare.2022.03.019
*Authors contributed equally to this work.

Britt M. Hermes participated in 16S rRNA gene data generation and curation, the formal analysis of the data, use of software to analyze and visualize data, defining study methodologies, data validation, and data interpretation for publication. Britt M. Hermes and Meriem Belheouane prepared the original manuscript; review and editing was supported by co-authors. Britt M. Hermes designed the graphical abstract. Britt M. Hermes was responsible for the planning, execution, and preparation of the work for publication.

CRediT authorship contribution statement:
Conceptualization: SI, CDS, ES, JFB; Data curation: BMH, MB; Study (experiment) conduct: MB, BMH; Formal analysis: BMH, MB; Funding acquisition: CDS, JFB, ES; Investigation: BMH, MB; Methodology: BMH, MB, JFB, CDS, ES; Project administration: NVB, CMH, MMH, ES, JFB; Resources: NVB, SB, PB, JFB, KB, SG, RG, MG, CG, AVG, CMH, MMH, BH, FH, BL, PJ, DM, AP, CP, MS, LH, SV, CDS, ES; Software: BMH, MB; Supervision: JFB, MB; Validation: BMH, JFB, MB; Visualization: BMH, MB, JFB; Writing–Original draft preparation: BMH, MB, JFB; Writing–Review and editing: all authors

Jan Schubert, Katja Cloppenborg-Schmidt, and Olga Eitel provided technical assistance. Sarah Gaugel and Stephanie Freyher, Lübeck, provided technical assistance with sample storage. Ana Luiza Lima, Kaan Yilmaz, and Onur Dikmen, Lübeck, helped with sample storage and communication with study centers during various phases of the study.

Chapter 2
Hermes, B. M., Rademacher, F., Chung, C. J., Tiegs, G., Bendix, M. C., de Zwaan, M., Harder, J., Baines, J. F. Skin microbiota analysis in patients with anorexia nervosa and healthy-weight controls reveals microbial indicators of healthy weight and associations with the antimicrobial peptide psoriasin. ***Scientific Reports*.** 12, 15515. (2022). https://doi.org/10.1038/s41598-022-19676-6

Britt M. Hermes participated in study conceptualization, 16S rRNA gene data generation and curation, the formal analysis of the data, use of software to analyze and visualize data, data interpretation for publication, defining study methodologies, project administration, project resources, and data validation. Britt M. Hermes prepared the original manuscript; review and editing was supported by co-authors. Britt M. Hermes was responsible for the planning, execution, and preparation of the work for resubmission with revisions and for publication. Britt M. Hermes is a corresponding author for this study.

CRediT authorship contribution statement:
Conceptualization: BMH, JH, MZ, GT, JFB, FR; Data curation: BMH (16S rRNA gene data); MB (AMP data); CJ (ddPCR data); Study (experiment) conduct: MB, MZ; Formal analysis: BMH; Funding acquisition: MZ, MB, JH, JFB; Investigation: BMH; Methodology: BMH, JH,

JFB; Project administration: BMH, JH, JFB; Resources: BMH, JH, JFB; Software: BMH; Supervision: JH, JFB; Validation: BMH, JFB; Visualization: BMH; Writing–Original draft preparation: BMH; Writing–Review and editing: all authors

Katja Cloppenborg-Schmidt and Yasmin Claussen provided technical assistance with laboratory procedures, including 16S rRNA gene amplicon library preparation.

Chapter 3
Hermes, B. M., Hirose, M., Tietje, A. M., Andreani, N. A., Belheouane, M., Ibrahim, S., Baines, J. F. Characterization of skin microbiota in a murine *Nell2* knockout model. Manuscript in preparation.

Britt M. Hermes participated in study conceptualization, collection of study material, 16S rRNA gene data generation and curation, the formal analysis of the data, use of software to analyze and visualize data, data interpretation for publication, defining study methodologies, project administration, project resources, and data validation. Britt Hermes prepared the original manuscript; review and editing was supported by co-authors. Britt M. Hermes was responsible for the planning, execution, and preparation of the work for publication.

CRediT authorship contribution statement:
Conceptualization: JFB, SI, BMH, MH, MB, NAA; Data curation: BMH (16S rRNA clone library, sequencing library, DNA/RNA extraction, cDNA, qPCR), MB (G15 AIL data), MH (*Nell2* sampling), AMT (*Nell2* sampling, *Nell2* genotyping), NAA (DNA/RNA extraction, cDNA, library preparation); Study (experiment) conduct: BMH, MH, AMT; Formal analysis: BMH, AMT (genotype data); Funding acquisition: JFB, SI; Investigation: BMH, AMT (genotyping); Methodology: BMH, JFB, SI, MH, AMT, NAA; Software: BMH; Resources: BMH, MH, JFB, SI; Supervision: JFB, SI, MH, NAA; Validation: BMH, JFB; Visualization: BMH; Writing–Original draft preparation: BMH; Writing–Review and editing: BMH, MH, SI, JFB

Katja Cloppenborg-Schmdit and Yasmin Claussen provided technical assistance with laboratory procedures, including qPCR, 16S rRNA gene amplicon library preparation, cloning, primer evaluation, and sanger sequencing. Dr. Malte C. Rühlemann and Dr. Shauni Doms provided statistical and analytical support.

Key vocabulary, abbreviations, and acronyms

Bacterial nomenclature
This thesis follows bacterial nomenclature guidelines set forth by the *American Society for Microbiology* and the *Journal of Bacteriology*, whereby all microbial taxa are italicized, strain designations and numbers are not. Binary names are used for all bacteria; names are abbreviated after the first use of a specific epithet (e.g., *Staphylococcus aureus* abbreviated thereafter as *S. aureus*).

Microbiome research terminology
This thesis follows vocabulary recommendations by Marchesi and Ravel (2015) to describe research methods and microbial community features and their associated environments. To preserve the standardization of accepted vocabulary in the field of microbiome research, I herein describe vocabulary, abbreviations, and acronyms used throughout this thesis:

Microbiota: the composition and abundance of microorganisms within a discrete environment, human, or otherwise. The term *microbiota* was first proposed by Lederberg and McCray (2001).

Microbiome: the collection of microorganisms, their genetic material, and their surrounding environment. This term is derived from the suffix "-biome," which includes both the biotic (living) and abiotic (non-living) elements of the given environment (Marchesi and Ravel, 2015).

Metagenome: the genetic material derived from microbiota. The collection of the metagenome is typically obtained through shotgun sequencing with subsequent assembly or mapping to reference databases (Marchesi and Ravel, 2015).

Metaorganism: a multicellular entity that comprises the interactions between a host species and its entire breadth of microbial communities (Turnbaugh et al., 2007; Bosch and McFall-Ngai, 2011).

16S rRNA gene: The 16S rRNA gene is named according to the following nomenclature: The "S" refers to the Svedberg unit, a non-SI unit for sedimentation rate. It is a measure of particle size relating to its rate of travel in a tube under high *g* force. 16S ribosomal ribonucleic acid (rRNA) refers to the ribosomal component of the body of the 30S subunit of bacterial and archaeal ribosomes. They make-up one structure of 16S rRNA that is bound to constituent 21 proteins. The genes that encode this structure are referred to as the 16S rRNA gene (Lederberg and McCray, 2001). Multiple 16S rRNA gene sequences can exist within a single bacterium (Case et al., 2007). The 16S rRNA gene is a popular molecular marker used in reconstructing phylogenies because of the slow rate of evolution within a conserved portion of this gene (Case et al., 2007; Woese and Fox, 1977).

16S rRNA gene sequencing/analysis: 16S rRNA gene sequencing is a common molecular method used to characterize mammalian resident microbiota. Analysis of 16S rRNA gene sequences includes the clustering of related sequences at a defined similarity threshold followed by counting the number of the representative sequences of each cluster (Byrd et al., 2018; Jo et al., 2016).

Operational taxonomical unit (OTU): An OTU refers to a cluster of closely related DNA sequences. In microbiome research, this is usually within the context of 16S rRNA gene sequencing and analysis, but it can also refer to other taxonomic marker genes. A threshold cut-off of 97% is the generally accepted similarity threshold (Kopylova et al., 2016; Stackebrandt and Goebel, 1994; Westcott and Schloss, 2015). In many cases, OTU is used as a proxy for "species"

(Blaxter et al., 2005).

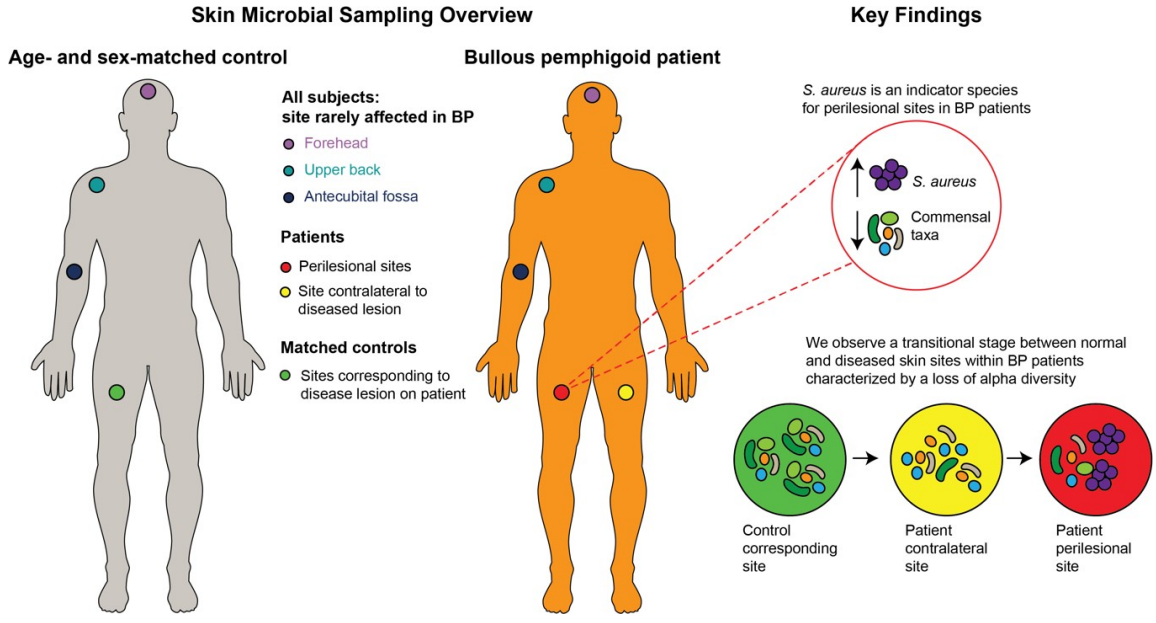
Amplicon sequence variant (ASV) or exact sequence variant (ESV): a single DNA sequence read derived from high-throughput marker gene (i.e., 16S rRNA gene) analysis. An ASV read is generated after the removal of erroneous sequences that are usually generated during PCR amplification and subsequent sequencing (Acinas et al., 2005; Kunin et al., 2010). Differing ASVs can vary by as few as one nucleotide; an ASV can be thought of as an OTU with 100% sequence similarity (Porter and Hajibabaei, 2018). Researchers argue ASVs offer a more precise and accurate measurement of sequence variation (Callahan et al., 2017), and therefore, some microbiome researchers have argued for defining taxa based solely on exact nucleotide sequences (Callahan et al., 2017). A recent study from 2019 concluded that ASV-based inference methods perform better than OTU clustering methods for distinguishing noise from biological signal in low biomass communities (Caruso et al., 2019). However, a slightly older study comparing binning approaches (i.e., OTU versus ASV) using a large field-based data set found that ASV binning provided just a minor improvement in taxonomic resolution over OTU binning, demonstrating that, in general, ribosomal genetic markers are inadequate molecular targets if high-genetic resolution of a microbial community is required (Glassman and Martiny, 2018). Moreover, OTUs are still widely used, especially for the comparison of diversity across large datasets, for example, (Delgado-Baquerizo et al., 2018).

Abbreviations and acronyms
AI: autoimmune
AIBD: autoimmune blistering disease
AMP: antimicrobial peptide
AN: anorexia nervosa
ASV: amplicon sequence variant
BP: bullous pemphigoid
BPDAI: Bullous Pemphigoid Disease Area Index
BMI: body mass index
DF: degrees of freedom
EBA: epidermolysis bullosa acquisita
GWAS: genome-wide association
IBD: inflammatory bowel disease
IS: indicator species
LPS: lipopolysaccharides
MR: Mendelian Randomization
OTU: operational taxonomic unit
PCR: polymerase chain reaction
SNP: single-nucleotide polymorphism
TLR: Toll-like receptor
QTL: quantitative trait locus analysis

Abstracts

Chapter 1

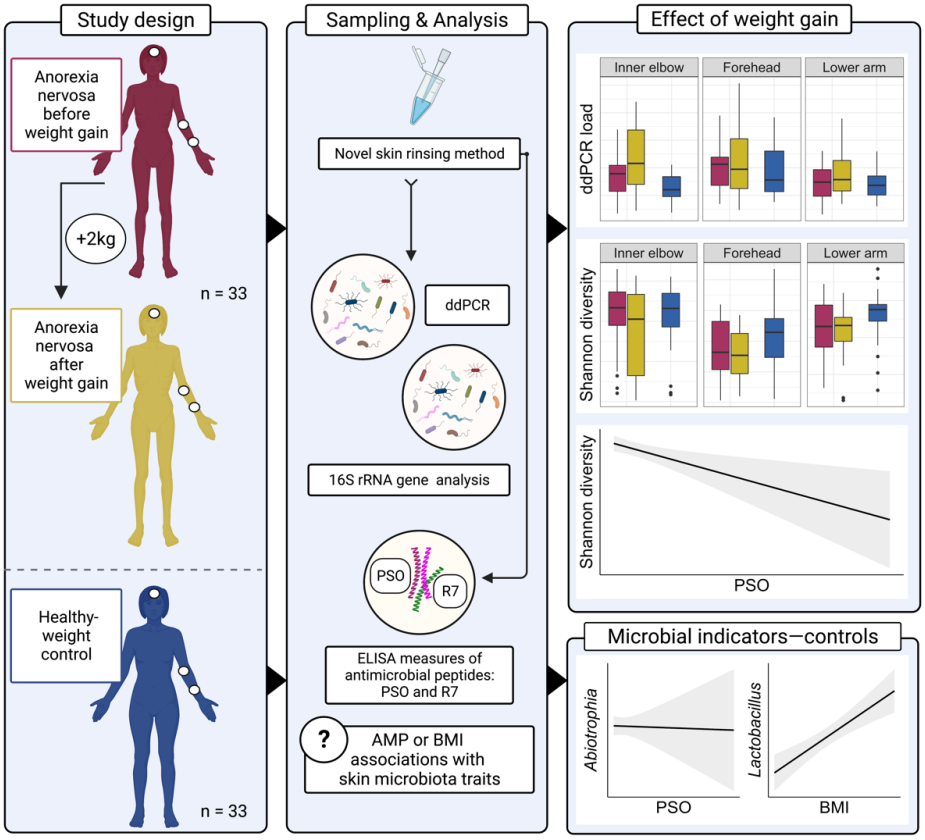
Bullous pemphigoid (BP) is the most common autoimmune blistering disease. It predominately afflicts the elderly and is significantly associated with increased mortality. The observation of age-dependent changes in the skin microbiota as well as its involvement in other inflammatory skin disorders suggests that skin microbiota may play a role in the emergence of BP blistering. We hypothesize that changes in microbial diversity associated with BP might occur before the emergence of disease lesions, and thus could represent an early indicator of blistering risk. The present study aims to investigate potential relationships between skin microbiota and BP and elaborate on important changes in microbial diversity associated with blistering in BP. This study consisted of an extensive sampling effort of the skin microbiota in patients with BP and age- and sex-matched controls to analyze whether intra-individual, body site, and/or geographical variation correlate with changes in skin microbial composition in BP and/or blistering status. We find significant differences in the skin microbiota of patients with BP compared to that of controls, and moreover that disease status rather than skin biogeography (body site) correlates with skin microbiota composition in patients with BP. Our data reveal a discernible transition between normal skin and the skin surrounding BP lesions, which is characterized by a loss of protective microbiota and an increase in sequences matching *Staphylococcus aureus,* a known inflammation-promoting species. Notably, *Staphylococcus aureus* is ubiquitously associated with BP disease status, regardless of the presence of blisters. Our findings suggest *Staphylococcus aureus* may be a key taxon associated with BP disease status. Importantly, we find that contrasting patterns in the relative abundances of *Staphylococcus hominis* and *Staphylococcus aureus* reliably discriminate between patients with BP and matched controls. This may serve as valuable information for assessing blistering risk and treatment outcomes in a clinical setting.



Graphical abstract designed by Britt M. Hermes Belheouane, Hermes, et al. 2022. *JARE*. https://doi.org/10.1016/j.jare.2022.03.019

Chapter 2

Anorexia nervosa (AN) is a psychiatric condition defined by low body weight for age and height and is associated with numerous dermatological conditions. Yet, clinical observations report that patients with AN do not suffer from infectious skin diseases such as those associated with primary malnutrition. Cell-mediated immunity appears to be amplified in AN, however, and this pro-inflammatory state does not sufficiently explain the lower incidence of infections. Antimicrobial peptides (AMPs) are important components of the innate immune system defending against pathogens and shaping the microbiota. In *Drosophila melanogaster*, starvation precedes increased AMP gene expression. Here, we analyzed skin microbiota in patients with AN and age-matched, healthy-weight controls and investigated the influence of weight gain on microbial community structure. We then correlated features of the skin microbial community with psoriasin and RNase 7, two highly abundant AMPs in human skin, to clarify whether an association between AMPs and skin microbiota exists and whether such a relationship might contribute to the resistance to cutaneous infections observed in AN. We find significant statistical correlations between Shannon diversity and the highly abundant skin AMP psoriasin and bacterial load, respectively. Moreover, we reveal that psoriasin significantly associates with *Abiotrophia*, an indicator for the healthy-weight control group. Additionally, we observe a significant correlation between an individual's body mass index and *Lactobacillus*, a microbial indicator in the healthy study group. Future investigation may help clarify physiological mechanisms that link nutritional intake with skin physiology.

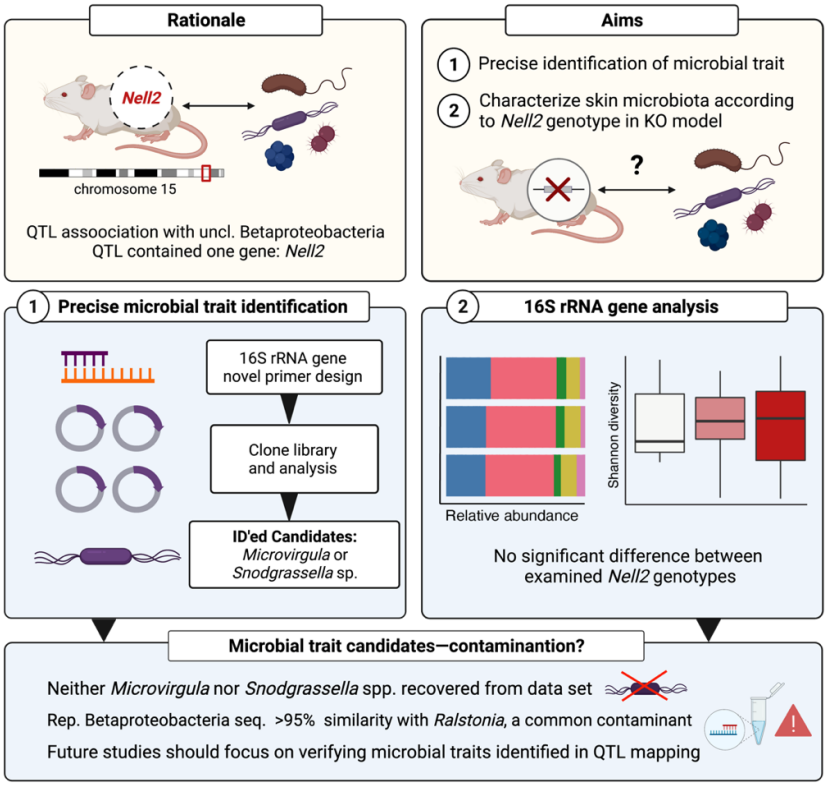


Graphical abstract designed by Britt M. Hermes.
Created with BioRender.com.

Chapter 3

Skin microbiota play a crucial role in skin biology, including moderating local inflammatory responses and immune cell functioning. Disruptions in the homeostasis between host and commensal skin microbiota may lead to chronic inflammatory skin diseases. Thus, characterizing the relationship between host genetics and the assembly of the skin microbiome is central to understanding how microbiota influence human health and whether microbiota could be exploited as therapeutic interventions. Previously, using the 15th generation of an advanced intercross line, we demonstrated that abundances of bacterial taxa in the skin might be significantly influenced by host genetic variation. One candidate region was associated with unclassified *Betaproteobacteria* and contained one gene: neural epidermal growth factor-like 2 (*Nell2*). *Nell2* is predominately expressed in neural tissues but has also been found to be differentially expressed in the epidermis of patients suffering from atopic dermatitis (AD).

While the relationship between *Nell2* and AD remains unelucidated, it is intriguing that an increased number of cutaneous free nerve endings has been observed in the epidermis of patients with AD, perhaps contributing to the intense pruritis that epitomizes this inflammatory skin disease. Here, we aimed to further explore the association between *Nell2* and unclassified *Betaproteobacteria* in more detail by precisely identifying the bacterial taxon involved and 16S rRNA gene amplicon sequencing and analysis of a *Nell2* knock-out strain. We reveal evidence suggesting that the unclassified *Betaproteobacteria* trait might instead belong to *Burkholderiaceae* within the class *Gammaproteobacteria*. Moreover, we find that unclassified *Betaproteobacteria* abundance does not significantly vary according to the examined *Nell2* genotype in the knock-out strain. We show that most features of the skin microbiota do not significantly differ between *Nell2* genotypes. We find evidence suggesting that the unclassified *Betaproteobacteria* trait might be a contaminant frequently found in DNA/RNA extraction kits. Our findings warrant future studies to validate host gene-microbe associations previously observed in genetic mapping studies involving murine skin.



Graphical abstract designed by Britt M. Hermes.
Created with BioRender.com.

Zusammenfassungen

Chapter 1
Das bullöse Pemphigoid (BP) ist die häufigste blasenbildende Autoimmunerkrankung. Sie betrifft vor allem ältere Menschen und führt zu einer deutlich erhöhten Sterblichkeitsrate. Altersabhängige Veränderungen der Hautmikrobiota sowie deren Beteiligung an anderen entzündlichen Hauterkrankungen legen nahe, dass die Mikroorganismen der Haut eine Rolle bei der Entstehung von BP-abhängiger Blasenbildung spielen könnten. Wir vermuten, dass Veränderungen der mikrobiellen Diversität im Zusammenhang mit BP vor dem Auftreten von Krankheitsläsionen auftreten und somit einen frühen Indikator für das Blasenbildungsrisiko darstellen könnten. Die vorliegende Studie zielt darauf ab, potenzielle Beziehungen zwischen der Mikrobiota der Haut und BP zu untersuchen und wichtige Veränderungen der mikrobiellen Diversität im Zusammenhang mit der Blasenbildung bei BP zu identifizieren. Um herauszufinden, ob intraindividuelle, körperliche und/oder geografische Unterschiede mit Veränderungen der mikrobiellen Zusammensetzung der Haut bei BP und/oder Blasenbildung korrelieren, wurden im Rahmen der Studie umfangreiche Proben der Hautmikrobiota von Patienten mit BP und alters- und geschlechtsgleichen Kontrollpersonen entnommen. Wir fanden signifikante Unterschiede der Hautmikrobiota von BP-Patienten im Vergleich zu der von Kontrollpersonen, und darüber hinaus, dass der Krankheitsstatus und nicht die Hautbiogeografie (Körperstelle) mit der Zusammensetzung der Hautmikrobiota bei BP-Patienten korreliert. Unsere Daten zeigen einen erkennbaren Übergang zwischen normaler Haut und der Haut, die BP-Läsionen umgibt. Dieser Übergang ist durch einen Verlust der schützenden Mikrobiota und einer Zunahme von Sequenzen gekennzeichnet, die zu *Staphylococcus aureus*, eines bekannten entzündungsfördernden Bakteriums, passen. Bemerkenswert ist, dass *Staphylococcus aureus* unabhängig vom Vorhandensein von Blasen ubiquitär mit dem Krankheitsstatus von BP assoziiert ist. Unsere Ergebnisse deuten darauf hin, dass *Staphylococcus aureus* ein Schlüsseltaxon, das mit dem BP-Krankheitsstatus assoziiert ist, sein könnte. Ein wichtiges Ergebnis der hier vorliegenden Studie ist, dass gegensätzliche Muster in den relativen Häufigkeiten von *Staphylococcus hominis* und *Staphylococcus aureus* zuverlässig zwischen Patienten mit BP und entsprechenden Kontrollen unterscheiden. Dies kann als wertvolle Information für die Beurteilung des Blasenbildungsrisikos und der Behandlungsergebnisse in einem klinischen Umfeld dienen.

Chapter 2
Anorexia nervosa (AN) ist eine psychiatrische Erkrankung, die durch ein im Verhältnis zu Alter und Körpergröße niedriges Körpergewicht gekennzeichnet ist und mit zahlreichen dermatologischen Erkrankungen einhergeht. Klinische Beobachtungen zeigen jedoch, dass Patienten mit AN nicht an infektiösen Hauterkrankungen leiden, wie sie zum Beispiel bei primärer Unterernährung auftreten. Allerdings scheint die zellvermittelte Immunität bei AN verstärkt zu sein. Dieser proinflammatorische Zustand erklärt die geringere Inzidenz von Infektionen jedoch nicht ausreichend. Ein wichtiger Teil des angeborenen Immunsystems sind Antimikrobielle Peptide (AMPs), die nicht nur gegen Krankheitserreger verteidigen, sondern auch die Mikrobiota formen. Bei *Drosophila melanogaster* geht Mangelernährung einer erhöhten AMP-Genexpression voraus. Hier haben wir die Mikrobiota der Haut von Patienten mit AN und altersgleichen, gesundgewichtigen Kontrollpersonen analysiert und den Einfluss der Gewichtszunahme auf die Struktur der mikrobiellen Gemeinschaft untersucht. Anschließend korrelierten wir Merkmale der Hautmikrobiota mit Psoriasin und RNase 7, zwei in der menschlichen Haut sehr häufig vorkommenden AMPs, um zu klären, ob ein Zusammenhang zwischen AMPs und der Hautmikrobiota besteht und ob eine solche Beziehung, zu der bei AN beobachteten Resistenz gegen Hautinfektionen beitragen könnte. Wir fanden signifikante statistische Korrelationen zwischen der Shannon-Diversität und dem in der Haut sehr häufig vorkommenden AMP Psoriasin bzw. der bakteriellen Belastung. Darüber hinaus zeigen wir, dass

Psoriasin signifikant mit Abiotrophia, einem Indikator für die gesundheitsbewusste Kontrollgruppe, assoziiert ist. Außerdem beobachten wir eine signifikante Korrelation zwischen dem Body-Mass-Index einer Person und *Lactobacillus*, einem mikrobiellen Indikator in der gesunden Studiengruppe. Künftige Untersuchungen könnten zur Klärung der physiologischen Mechanismen beitragen, die die Nahrungsaufnahme mit der Hautphysiologie verbinden.

Chapter 3

Die Hautmikrobiota spielt eine entscheidende Rolle in der Hautbiologie, u. a. bei der Regulierung lokaler Entzündungsreaktionen und der Funktion von Immunzellen. Störungen der Homöostase zwischen Wirt und Hautmikrobiota können zu chronisch entzündlichen Hautkrankheiten führen. Um zu verstehen, wie die Hautmikrobiota die menschliche Gesundheit beeinflusst und ob diese als therapeutische Maßnahme genutzt werden könnte, ist die Charakterisierung der Beziehung zwischen der Genetik des Wirts und dem Aufbau des Hautmikrobioms von zentraler Bedeutung. Anhand der 15. Generation einer fortgeschrittenen Kreuzungslinie konnten wir nachweisen, dass die Häufigkeit von Bakterientaxa in der Haut erheblich von der genetischen Variation des Wirts beeinflusst werden kann. Eine außergewöhnliche Kandidatenregion war mit nicht klassifizierten *Betaproteobakterien* assoziiert und enthielt das Gen *Nell2* (neural epidermal growth factor-like 2). *Nell2* wird vorwiegend in Nervengeweben exprimiert. In Patienten mit atopischer Dermatitis (AD), einer entzündlichen Hauterkrankung, die durch starken Juckreiz charakterisiert wird, wird *Nell2* zusätzlich auch in der Epidermis differentiell exprimiert.

Obwohl der Zusammenhang zwischen *Nell2* und AD noch nicht geklärt ist, ist es interessant, dass in der Epidermis von AD-Patienten eine erhöhte Anzahl von freien Nervenenden beobachtet wurde, die möglicherweise zu dem starken Juckreiz beitragen. In der hier vorliegenden Studie wollten wir die Assoziation zwischen *Nell2* und nicht klassifizierten *Betaproteobakterien*, durch eine genaue Identifizierung des beteiligten Bakterientaxons mittels 16S rRNA-Sequenzierung und die Analyse eines *Nell2*-Knock-out-Stamms näher untersuchen. Wir finden Hinweise darauf, dass die nicht klassifizierten *Betaproteobacteria* zu *Burkholderiaceae* innerhalb der Klasse *Gammaproteobacteria* gehören könnten. Außerdem fanden wir heraus, dass sich die Häufigkeit von nicht klassifizierten *Betaproteobacteria* nicht signifikant zwischen dem *Nell2*-Knock-out-Stamms und des Kontrollstamms unterscheiden. Zusätzlich zeigen wird, dass die Merkmale der Hautmikrobiota zwischen den *Nell2*-Genotypen nicht signifikant unterschiedlich sind. Schließlich finden wir Hinweise darauf, dass die nicht klassifizierten *Betaproteobacteria* eine häufige Verunreinigung in DNA/RNA-Extraktionskits sind. Unsere Ergebnisse rechtfertigen zukünftige Studien, die darauf abzielen, Assoziationen zwischen Wirtsgenen und Bakterien zu validieren, die zuvor in genetischen Kartierungsstudien mit Mäusehaut beobachtet wurden.

Introduction

The skin microbiome, our hidden organ on the surface

The skin is the largest organ in the human body, approximately 1.5 to 2 m^2 in area, and serves as a crucial physical and immune barrier from our environment. Yet, the skin is home to billions of microorganisms inhabiting a multitude of folds, invaginations, and specialized niches that sustain microscopic life (Grice and Segre, 2011; Scharschmidt and Fischbach, 2013). We are not teeming with bugs so much as we are irrevocably intertwined with microbiota such that, throughout evolutionary history, these bacteria, fungi, and viruses have become invaluable parts of our physiology, cooperatively acting as a hidden organ. The microbiota colonizing the skin, their genetic material, and their microenvironments are collectively referred to as the skin microbiome (Marchesi and Ravel, 2015).

Most microorganisms are not pathogens. Rather, the majority of microorganisms living on their hosts are benign and, in some cases, beneficial, performing functions vital for host physiology and homeostasis. Scientists are beginning to unravel how the skin microbiome is interfacing with different physiological processes of the mammalian host. Microbiota living on us, but also residing within deeper layers of the skin, contribute to host inflammation, epidermal barrier function, and immunity through feedback mechanisms with our immune cells (Grice and Segre, 2011). Microbiota even contribute to host defense through the production of bactericidal factors (Iacob et al., 2018; Jacobs et al., 2017; Nakatsuji et al., 2017; Zheng et al., 2020). Microbiome researchers are interested in understanding these mechanisms that link commensal microbes with host biology and how these interactions contribute to host health or disease. This thesis aims to characterize resident skin microbiota communities the context of inflammatory cutaneous disease.

Human skin: an overview

Human skin comprises the epidermis and the dermis layers, separated by a basement membrane. Sweat glands, hair follicles, sensory neurons, and blood vessels are housed within the deep layer of the dermis. The epidermis, the uppermost layer, provides a defense against invaders as an adaptive immunological barrier, a chemical/biochemical barrier, and a physical barrier. Humoral and cellular components of the immune system compose the adaptive immunological defense system (Proksch et al., 2008). The chemical barrier is formed by components of innate immunity, specifically lipids, acids, enzymes, antimicrobial peptides, and macrophages. The physical barrier is composed of the stratum corneum (the outmost layer), which protects against physical, chemical, and microbial incursions on the host. Collectively, these physiological components of the epidermis serve as formidable armor against pathogens and environmental attacks, while simultaneously retaining water and nutrients essential for host health (Grice and Segre, 2011).

Remarkably, the skin is a self-renewing organ capable of replenishing itself every few weeks (Fuchs and Raghavan, 2002). The stratum corneum is composed of dead, flattened, enucleated keratinocytes known as squames (Fuchs and Raghavan, 2002; Grice and Segre, 2011). Squames are continuously sloughed off and replaced by inner keratinocytes that have detached from the cell cycle in the basement membrane as a final step in terminal differentiation and moved outwards toward the stratum corneum. These dead cells comprise a tough, brick and mortar-like exterior that tends to be dry, cool, and acidic (Belkaid and Segre, 2014; Grice and Segre, 2011). Interestingly, human skin appears to impart selective pressures on its microbes (Vandegrift et al., 2017). Human skin is in constant contact with the environment, yet it harbors comparatively low-biomass communities.

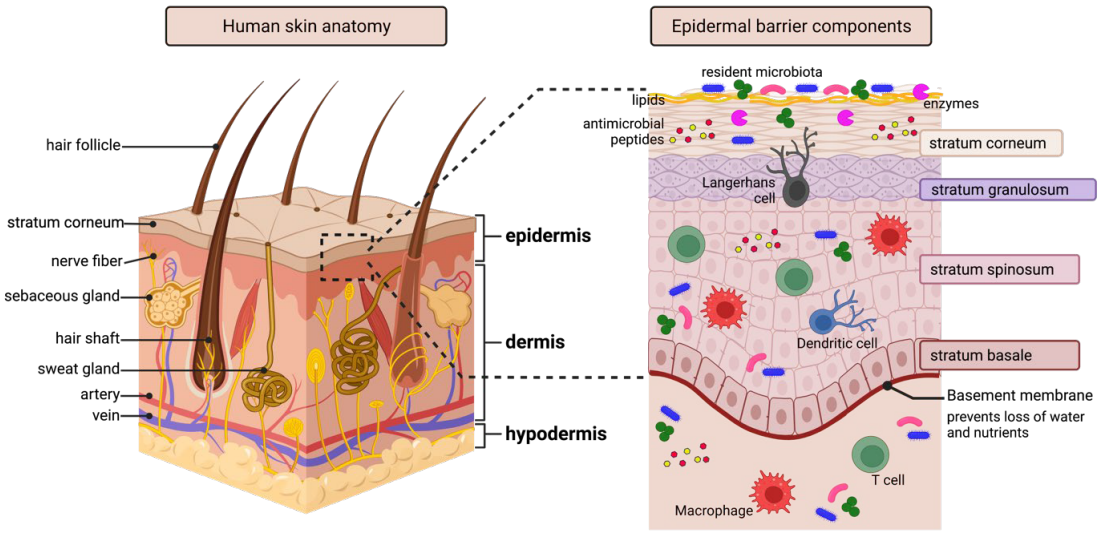


Figure 1. Human skin anatomy. Human skin consists of the epidermis and dermis, separated by an extracellular matrix, called the basement membrane, that provides a stabile foundation for the epidermis. It additionally helps prevent loss of water and nutrients. The epidermidis consists of physical, chemical, and immunological components that collectively serve as a protective barrier against microbial and environmental assaults on the host. Langerhans cells, dendritic cells, T cells, antimicrobial peptides, and enzymes reside within the skin. Resident microbiota live on and within the deeper layers of healthy human skin and contribute to host defense. The stratum corneum, is a dry, cool, and acidic exterior comprised of squames. Epidermal surface lipids, e.g., ceramides, free fatty acids, and cholesterol, help prevent water and electrolyte loss in additional to protective functions. Image by Britt M. Hermes. Created with Biorender.com.

Human skin biogeography

The topographical variation of the skin's surface, which includes numerous topographic features and specialized appendages such as sweat glands and hair follicles, provides distinct environments suitable for microbial life (Costello et al., 2009; Fuchs and Raghavan, 2002; Grice et al., 2009; Grice and Segre, 2011). Accordingly, human skin displays "biogeography," whereby distinct microbial assemblages colonize body sites depending upon temperature, moisture, pH, antimicrobial peptides, lipids, and physical skin structures (Costello et al., 2009; Grice et al., 2009; Oh et al., 2014). In other words, the three major skin habitats, i.e., dry, moist, and sebaceous, relate to the structure and make-up of the bacterial communities living there.

The skin microbiome is composed of a limited number of bacteria, most of which are Gram-positive species. The scientific consensus is that species from *Staphylococcus*, *Corynebacterium*, *Propionibacterium*, *Micrococcus*, *Streptococcus*, *Brevibacterium*, *Acinetobacterium*, and *Pseudomonas* genera are the most abundant commensal human skin residents (Byrd et al., 2018; Grice and Segre, 2011; Zhou et al., 2020). The sebaceous, oily zones, such as the upper back, forehead, and chest, tend to be low in microbial diversity and dominated by lipophilic microorganisms such as *Propionibacterium acnes* and *Malassezia,* a fungus (Grice et al., 2009; Grice and Segre, 2011). *Betaproteobacteria*, from the phylum *Proteobacteria*, also inhabit oily skin sites (Grice and Segre, 2011). Moist body zones, including the groin, armpit, and toe webs, support microorganisms that thrive in humid conditions, such as Gram-negative bacilli, coryneforms, and *Staphylococcus aureus* (Grice et al., 2009). Dry zones, such as the forearm, palm of hand, and leg, are often the richest in bacterial diversity (Grice et al., 2009). Both dry and moist zones tend to be more vulnerable to temporal variability, picking-up bacterial "hitch-hikers" that seem only to colonize the skin transiently (Grice et al., 2009; Oh et al., 2016).

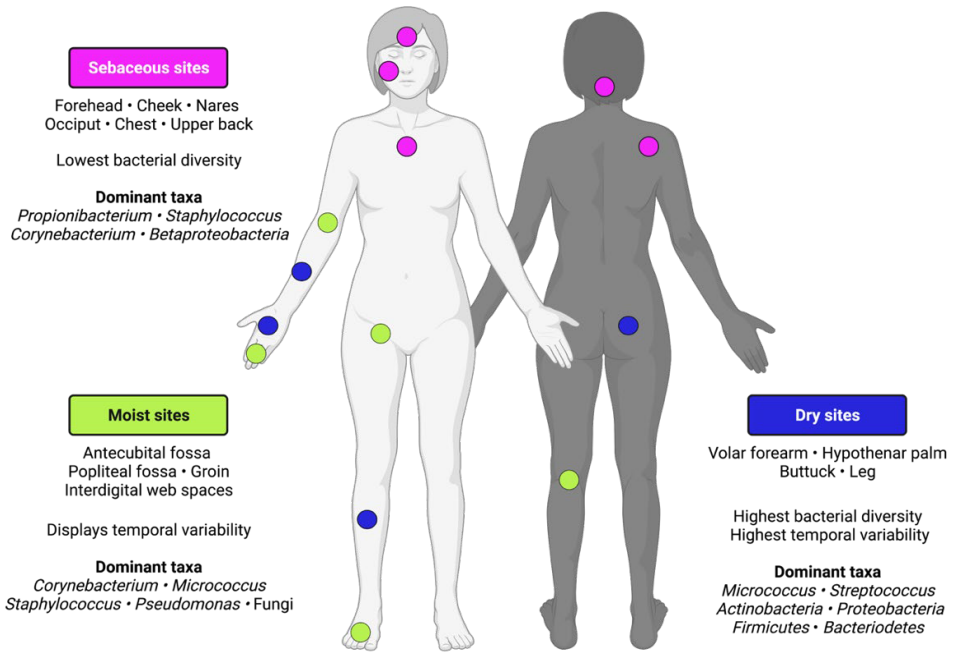


Figure 2. Human skin biogeography. Humans are colonized by microbial assemblages that relate to sebaceous, moist, or dry body sites. These body sites vary in temperature, moisture, pH, antimicrobial peptides concentrations, surface lipid composition, and physical skin structures. Image by Britt M. Hermes. Created with Biorender.com.

Distinct skin sites inhabit representative microbiota communities with abundant bacterial residents that are stable between individuals and over time (Costello et al., 2009). For example, Grice and colleagues (2011) found that body sites, such as the armpit, nares, back, and plantar heel, are more analogous in microbiota to the same body site on different individuals compared to that on a different body site of the same individual. Additionally, in a biogeographical survey of bacterial community variation, Costello and colleagues (2009) found that microbial communities tended to cluster according to body habitat, rather than by host sex, individual, or sampling time. For instance, the head region, including the external ears, forehead, hair, and external nose, could be differentiated from other body sites, including the arm, trunk, and leg. Further, UniFrac distance, which is a phylogeny-based measure used to assess bacterial community diversity, revealed that compositional variation in skin bacterial communities was higher between different body sites within individuals compared to similar body sites between individuals. Some sites on the body are alike, but these places tend to share ecological landscapes, such as the antecubital fossa (armpit) and popliteal fossa (back of knee) (Costello et al., 2009; Grice et al., 2009; Grice and Segre, 2011). In general, symmetric body sites within individuals tend to be more similar to each other compared with similar body sites between individuals (Costello et al., 2009; Grice et al., 2009, 2008; Grice and Segre, 2011).

Costello and colleagues (2009) aimed to disentangle whether biogeographical patterns observed at skin sites were determined by site-specific microenvironmental factors (*e.g.,* nutrient availability or local biochemical factors related to host ecological body niche), the availability of foreign microbes to colonize the site, or both. Using areas on the forehead and the volar forearm, the authors disinfected and then inoculated the chosen area with tongue microbiota; they subsequently tracked community changes over time. The authors found that at two, four, and eight hours after exposure to tongue microbes, the forearm bacterial community was more similar to the tongue rather than to the native forearm community. Conversely, the forehead became increasingly more similar to its native community over this short time span. The authors

concluded that at sebaceous sites such as the forehead, compared to dry sites like the forearm—which are generally much higher in bacterial diversity—the microenvironmental niche of the skin plays a pivotal role in shaping bacterial communities. These results were later reaffirmed in 2011 by Grice and Segre, concluding that ecological body site niche is a strong determinant of the microbiota composition in healthy individuals.

The dynamic human skin—resident bacteria ecosystem

Human skin bacteria are commonly categorized into two groups: resident and transient. Resident bacteria represent those species that are stable on human skin over time and are considered difficult to eradicate. Transient bacteria are assumed to be acquired by contact and easy to remove. In general, resident bacteria are considered commensal, i.e., neutral or beneficial for the host, and transient bacteria are considered contaminants, i.e., potentially pathogenic, or harmful, for the host. However, this terminology is misleading and over-simplifies the dynamism of the human skin-resident microbe ecosystem (Cogen et al., 2008).

For example, *Staphylococcus epidermidis* has long been considered a resident species that benignly benefits from receiving protection and nutrients from its specialized niche on human skin (Cogen et al., 2008; Vandegrift et al., 2017). But evidence accumulated over the past decade illustrates how this species can interact with human physiological processes along a mutualist–pathogen continuum. Nakatsuji and colleagues (2017) have demonstrated that *S. epidermidis* plays an active role in host defense through the production of bactericidal factors, such as antimicrobial peptides—a relationship representing mutualism (Cogen et al., 2008). However, *S. epidermidis* is also a well-documented opportunistic pathogen; it is responsible for the majority of infections that were acquired by patients with in-dwelling and implanted medical devices, such as catheters and replacement heart valves, respectively (Uçkay et al., 2009). In this regard, Cogen and colleagues (2008) specify that it is the host's capacity to resist microbial infection, rather than any bacterium's inherent pathogenicity, that discriminates benign from harmful bacteria. This includes, in part, host skin integrity and its epidermal biochemical and adaptive immunological barrier functions. Of note, *S. epidermidis* is routinely detected in the dermis of human skin where it can interact with host cells below the epidermal basement membrane (Nakatsuji et al., 2013). There are likely many members of the "healthy" skin microbiome that are associated with disease under opportunistic conditions.

Conversely, there are likely many so-called transient or pathogenic microbes that could be important to the structure of skin microbiota. *Staphylococcus aureus*, which is widely considered a pathogenic species and associated with atopic dermatitis, an inflammatory skin disorder, is commonly found on healthy skin, particularly in the nares (Peacock et al., 2001). One study from Poland found that up to 80% of healthy individuals were colonized with *S. aureus* (Masiuk et al., 2021). Another found that *S. aureus* colonized the nares of about 30% of the healthy population surveyed from the US (Graham et al., 2006). Cogen and colleagues (2008) summarize the bottom-line of these dynamic relationships aptly: "it is important to recognize that resident does not necessarily equate to commensal nor does transient always mean pathogenic."

Qualitative studies agree that skin microbiota are predominately comprised of a handful of stable bacterial inhabitants, i.e., *Propionibacterium* and *Staphylococcus* spp. While there are predictable bacterial community structures in similar body sites between individuals, enough rare and/or transient species are present to impart individual variation, making it difficult to define "core microbiota" of the skin (Oh et al., 2014; Vandegrift et al., 2017). The factors driving this variability are still unclear. Large, descriptive studies will be required to statistically determine the bacterial species that reliably colonize particular body sites and those species that are unique to individuals. However, an individual's microbial fingerprint tends to be remarkably stable over time despite an ever-changing environment (Costello et al., 2009; Grice et al., 2009; Oh et al.,

2016). This suggests that in addition to other principles governing community assembly and stability, host genetics likely contribute to the make-up of our skin microbiome.

Forces shaping host-associated microbiota: genes versus environment

Considerable effort has been invested into understanding the fundamental factors underlying how host-associated bacterial communities become established and are maintained over time. Broadly speaking, studies have identified diet, environment, and host genetics as factors that likely contribute to the make-up of our microbiota (Blekhman et al., 2015; De Filippo et al., 2010; Goodrich et al., 2016, 2014; Org et al., 2015; Song et al., 2013; Spor et al., 2011; Wang et al., 2016; Wu et al., 2011; Xu and Knight, 2015). These studies are largely limited to the gut. Nevertheless, such gut microbiome studies serve as models for studying skin-associated microbiota.

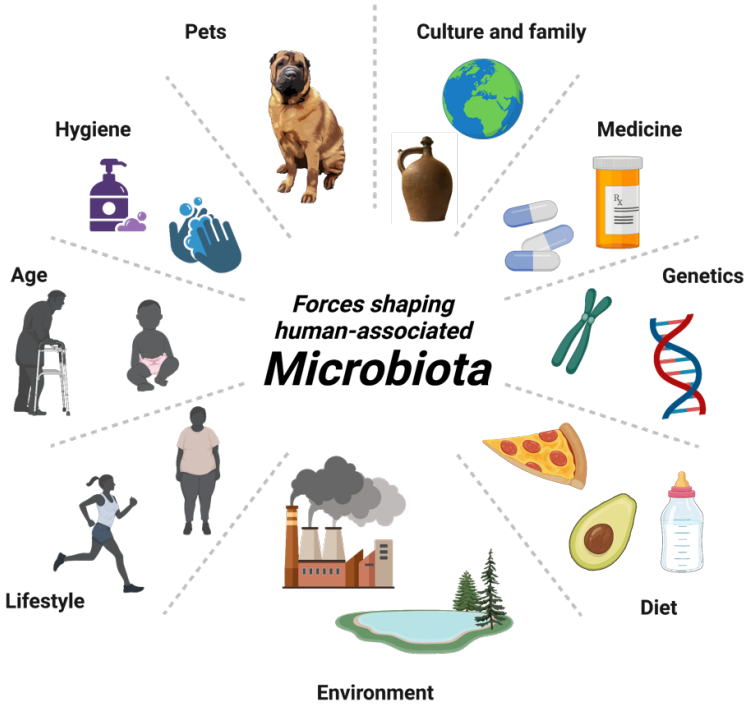


Figure 3. Forces shaping human-associated microbiota. Numerous factors, including diet, environment, pet ownership, and genetics contribute to the make-up of human-associated microbial communities. Image by Britt M Hermes. Created with Biorender.com.

Studies evaluating the role of genetic versus environmental factors in the assembly of the gut microbiome provide insight into the influence genes might have on host-associated microbiota community structure at other body sites. Research suggests there is a substantial genetic component of the host underlying interpersonal variability in gut microbiota, as evidenced by twin studies (Goodrich et al., 2016, 2014; Turnbaugh and Gordon, 2009; Yatsunenko et al., 2012) and the detailed analysis of individual candidate genes (Blekhman et al., 2015; Goodrich et al., 2016). Genetically related individuals tend to harbor more similar gut bacteria regardless of whether the individuals are cohabitating (Stewart et al., 2005; Yatsunenko et al., 2012; Zoetendal et al., 2008), and moreover, increased genetic similarities relates to increased similarities in gut community structure, as evidenced by detailed analyses comparing monozygotic and dizygotic twins (Goodrich et al., 2014). Twin designs are advantageous in that they allow researchers to

estimate genetic versus environmental effects on features of the microbiome. However, they are unable to identify specific environmental features or genes that contribute to phenotypic variance (Røysamb and Tambs, 2016).
Genome-wide association studies (GWAS) offer an alternative method to assess the role of genetics in shaping the assembly of host-associated microbiota. This observational study approach surveys the genomes of study participants for statistical associations with microbial traits, allowing researchers to identify host genes that might affect microbial community structure (Awany et al., 2019). Moreover, GWAS can be combined with other methods, such as Mendelian Randomization (MR), to infer causality between host-genes and microbes (Smith and Ebrahim, 2008). GWAS are challenging to perform because they require large sample sizes to achieve statistical power (Klein, 2007; Spencer et al., 2009; Awany et al., 2019). However, the curation of large biobanks allows researchers to perform well-powered genome-wide association analyses (Awany et al., 2019).

To further elucidate the role of host-genetics in shaping patterns of inter-individual variation in gut microbiota, we conducted a large-scale GWAS comprising 8,956 German individuals derived from five independent cohorts from German biobanks with subsequent MR analysis (Rühlemann et al., 2021). Our analyses identified 44 genome-wide significant associations with microbial features, including community composition. These associations comprised 32 genomic loci, some of which suggest immune-mediated interactions between the host and commensal microbiota (Rühlemann et al., 2021). For example, an association emerged between *Barnesiella* species and variants within the gene locus for biliverdin reductase A (*BLVRA*), which has been shown to inhibit gene expression for the pattern recognition receptor toll-like receptor 4 (*TLR4*) (Wegiel et al., 2011). These highly conserved receptors recognize pathogen-associated molecular patterns and are integral to the first-line defense system for protecting the host against microbial invaders (Molteni et al., 2016). *TLR4* recognizes the receptor for gram-negative bacterial lipopolysaccharide (LPS) (Paik et al., 2003; Molteni et al., 2016). Interestingly, *Barnesiella* is a gram-negative bacterium (Ormerod et al., 2016). The identified association suggests that *Barnesiella* might be one of the many microbiota that contribute to establishing immune tolerance of intestinal commensal microbiota early in life (Gensollen et al., 2016; Zheng et al., 2020).

Our GWAS additionally revealed a variant in the histo-blood group ABO system transferase (*ABO*) gene associates with an increased prevalence of *Bacteroides* operational taxonomic unit (OTU) 97_27. This *Bacteroides* OTU also significantly associates with variants at the BTB domain and CNC homolog 2 (*BACH2*) locus. Moreover, we find a non-significant correlation between this same *Bacteroides* OTU and variation at the galactoside 2-α-l- fucosyltransferase 2 (*FUT2*) locus (Rühlemann et al., 2021). Previous studies have reported that variants in the *BACH2* and *FUT2* genes are associated with inflammatory bowel disease (IBD) (Klasić et al., 2018; de Lange et al., 2017; Liu et al., 2015; McGovern et al., 2010; Momozawa and Dmitrieva, 2018). MR analysis suggests that *Bacteroides* OTU97_27 significantly protects against the development of Crohn's disease, a major form of IBD. This finding contrasts with previous surveys showing that lower abundances of *Bacteroides* species associate with the development of IBD (Nomura et al., 2003; Zhou and Zhi, 2016; Zuo and Ng, 2018). These data suggest interpersonal gut microbiome variation is shaped by host genetics, but how these host gene-microbe associations relate to disease susceptibility is still poorly understood.

Genetics are not the sole driver of microbiota community composition. In a study investigating how the gut microbiome might change over time within geographically distinct human populations, Yatsunenko and colleagues (2012) observed that the fecal microbiota of American teenagers were more like their biological parents than unrelated adults, and equally similar to that of their father and mother. But interestingly, the authors also found that mothers and fathers shared more gut bacterial taxa compared with individuals from other families. These results

suggest that shared environment and lifestyle factors (e.g., diet, hygiene practices, and culture) also significantly contribute to the fundamentals of gut microbial similarity amongst individuals living together.
Likewise, environmental features impact the skin microbiome of cohabitating individuals. Song and colleagues (2013) observed a strong familial membership effect on the variation and composition of human skin microbiota across all sampled body sites. They further observed that the effect of a shared environment was greater for skin microbiota compared to fecal and oral microbiota. Interestingly, features such as dog ownership impart a significant effect on the structure of skin microbiota. After controlling for age, Song and colleagues observed that adults who own dogs but do not live together, share, on average, as many bacterial OTUs as adults who cohabitate. Further, adults who own dogs tend to have greater bacterial diversity on their skin, likely due to the repeated introduction of rare bacteria from the environment into the household (Song et al., 2013).

These results are not surprising. The skin is in constant contact with its environment. Some body sites, like the dry areas of the forearm or leg, seem particularly susceptible to colonization by environmental microbes, some of which are left behind by pets or household members. Even household objects and the household air we breathe are largely dominated by skin-associated microbes (Fierer et al., 2010; Song et al., 2013). Therefore, it seems reasonable that resident skin communities would reflect the microbial communities in our local environments.

What is surprising, though, is that our skin microbial communities are largely stable at the strain level over time, despite constant environmental changes and exposure to other individuals (Oh et al., 2016). Song and colleagues (2013) observed that the skin microbiota of children aged six months to 18 years are relatively similar to that of genetically related adults in the same household. This contrasts with gut microbiota, which has a well-documented relationship to age and is known to change substantially between birth and three years of age, even among children in the same household as genetically related adults who are exposed to similar diets and living conditions (Koenig et al., 2011; Palmer et al., 2007; Yatsunenko et al., 2012). Further, in a longitudinal study of twelve individuals testing the stability of skin microbiota over time, Oh and colleagues (2016) found that both short-term (one to two months) and long-term (one to two years) community similarity was significantly greater within an individual compared to interindividual similarity. Notably, this study demonstrated diversity at skin sites was inversely correlated with temporal stability, i.e., increased diversity at skin sites relates to decreased stability over time. However, the community structure of known highly diverse sites, like dry zones and sites frequently exposed to the environment (*e.g.,* palm of hand) were still significantly stable over the observed time frames. Collectively, these data suggest that skin bacterial communities are generally homeostatic in healthy individuals, and that interactions between host and resident microbiota likely impact the structure of the skin microbiome.

The role of host genetic variation in the assembly of the skin microbiome was previously demonstrated in mice using a QTL mapping approach in a fourth-generation advanced intercross line (AIL) (Srinivas et al., 2013). Here, the authors reported thirteen regions in the mouse genome that were significantly associated with skin bacterial traits. However, the sizes of the loci (9 to 33 megabases) were too large to identify candidate genes. A follow-up study conducted by Belheouane and colleagues (2017) employed the fifteenth generation of this AIL, which dramatically increased the marker density of informative SNPs to improve the mapping resolution of QTLs for skin microbial abundances. The authors profiled 16S rRNA amplicon sequences at both the DNA and RNA levels, which reflect relative bacterial cell number and activity, respectively, and they identified numerous genomic loci associated with skin microbial abundances, with some containing single immune- and/or skin cancer-related genes. This study using high resolution mapping and phenotyping based on microbial activity underscores the

importance of elucidating the genetic component in interindividual variability of the skin microbiome. Characterizing the relationship between host genetics and the assembly of the skin microbiome is central for understanding how microbiota influence human health and whether microbiota could be exploited in therapeutic interventions (Ley et al., 2006; Turnbaugh et al., 2007).

Studying the skin microbiome: key considerations

The skin microbiome presents unique challenges in the field of microbiome research. As the skin harbors relatively low microbial biomass, the risk of contamination during skin sampling and sample processing is substantial and any contamination introduced during these steps can radically affect data interpretation, as contaminants tend to be preferentially amplified and sequenced over true biological signals within the sample (Eisenhofer et al., 2019; Karstens et al., 2019; Salter et al., 2014). Moreover, the range of diverse microenvironments (i.e., dry, moist, sebaceous) encompassing human skin as well as the need to consider bacteria living on the skin's surface and those residing within its deeper layers necessitate careful planning and consideration of potential pitfalls when conducting a skin microbiome study (Grice et al., 2008; Kim, 2017; Kong, 2016; Kong et al., 2017; Meisel et al., 2016; Nakatsuji et al., 2013). Therefore, skin microbiome research requires technical competence and skill in and out of the lab for reliability and reproducibility (Dahlberg et al., 2019; Eisenhofer et al., 2019; Glassing et al., 2016; Karstens et al., 2019; Salter et al., 2014). Here, I summarize study components critical for reliable skin microbiome research.

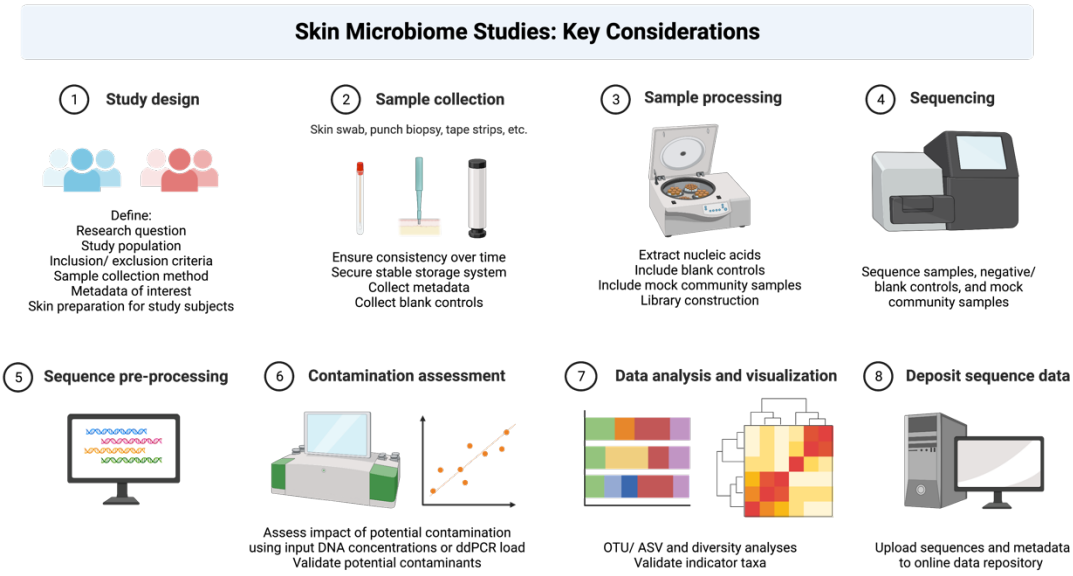


Figure 4. Key considerations for planning a skin microbiome study. The steps requiring careful consideration when planning a skin microbiome study include study design, sample collection and processing, sequencing, pre-processing of data, assessing for potential contaminants, and data analysis and visualization. Depositing sequencing data into an online data repository contributes to study reliability and replicability. Image by Britt M. Hermes. Created with Biorender.com.

Study design

Study design needs to be prudently planned prior to sample collection. The study population should be defined in advance and considered with regards to age, geographical location, disease state, and local environmental features, all of which are significantly associated with skin microbial diversity (Byrd et al., 2018; Grice and Segre, 2011; Gupta et al., 2017; Huang et al., 2020; Song et al., 2013; Yatsunenko et al., 2012). Detailed patient histories should be provided in study metadata, including medical history, medications, use of topical products, age, weight, and skin care regimens. It is essential to establish clinical protocols with detailed inclusion and

exclusion criteria. Specific instructions for participants to abstain from use of topical therapies, products, or other activities, such as exercise, should be carefully documented. Of particular importance is antibiotic use prior to sample collection, even within healthy individuals. The exclusion criteria for prior antibiotic use have been previously defined as 12 months, 6 months, and 1 month (Kong et al., 2017). When studying participants with a particular skin disease, clinical details are often critical for downstream analyses. Disease severity scores as well as disease phenotyping can be essential for understanding disease progression or subpopulations within a particular disease.

For example, in a large-scale investigation of patients with bullous pemphigoid (BP), an AI skin blistering disease, we observed a marked reduction in alpha diversity at blistering sites and at the body sites contralateral to blistering sites (Belheouane et al., 2022). Furthermore, we identified contrasting patterns of *Staphylococcus* ASVs in patients with BP compared to patients with non-inflammatory skin diseases. When we accounted for BP disease severity, as defined by an active disease score, we were able to differentiate the association of disease versus the severity of disease with skin microbiota within our study population. Surprisingly, active disease severity scores did not correlate with mean alpha diversity measures at diseased sites or at body sites contralateral to disease blisters/ erosions, suggesting that BP disease at any stage is significantly associated with reduced alpha diversity. These results provide further insight into the role of the skin microbiome in the emergence of blistering in BP and suggest that changes in the skin microbiota might occur early in the progression of disease. Such findings may be shown to have clinical relevance in future investigations (Belheouane et al., 2022).

Sampling methods and processing

Laboratory protocols should be established, and enough reagents and materials ordered in advance to ensure consistent methodologies over the course of the study. This can be a particular challenge in longitudinal studies, where suppliers might undergo shortages of lab materials or reagents, due to, as one example, a global pandemic affecting material production and supply (Woolston, 2021).

Skin sample collection methods previously employed in microbiome studies include swabbing, biopsies, surface scraping, cup scrubs, and tape stripes (Kong et al., 2017). The optimal strategy largely depends on the study question and design. Each method carries benefits and drawbacks, including patient discomfort, sampling depth, biomass yield, and the amount of host DNA potentially occluding bacterial DNA recovery (Kong et al., 2017). The sampling methodology should remain consistent throughout the study. These procedures include the time of day, sampling room, the person collecting the sample, as well as consistent use of buffer solutions, swabs, and other materials used during collection. Moreover, limiting the number of potential confounding variables, such as the number of individuals handling and processing the samples and the local sampling environment, can make a significant impact in downstream analyses (Kong et al., 2017).

Sample storage is a crucial, yet perhaps underappreciated, aspect of robust skin microbiome research. In general, processing of fresh samples is ideal, as this limits bias and degradation of bacterial DNA and/or RNA (Kong et al., 2017; Lauber et al., 2010). However, it is not always possible to process samples as they are collected, especially in the context of longitudinal studies and/or studies involving thousands of samples aimed at characterizing microbiota over space and time. Therefore, reliable storage strategies are required to ensure sample preservation and reproducible results. Ultimately, the best storage strategy will depend on the sampling method utilized in the study.

Mitigating contamination

There are two main types of contamination that can occur in microbiome studies. The first is cross-contamination from other samples and sequencing runs. Cross-contamination arises from the transfer of a DNA sample, barcode sequence, or amplicon from one well or tube into another, usually neighboring, well or tube during sample processing (Eisenhofer et al., 2019). This can happen throughout sample processing at several different steps. For example, cross-contamination occurs with pipetting errors, with accidental placement of a sample into the wrong tube/ well, or with aerosolization during pipetting and/or PCR plate covering/ removing. "Tag switching" represents another kind of cross-contamination that results when samples barcoded with molecular identifiers accidentally cross over, or "jump," into neighboring wells/tubes (Carlsen et al., 2012; Weyrich et al., 2019). This phenomenon can be controlled for by tagging both ends of the PCR amplicon, allowing researchers to identify non-compatible tag combinations after sequencing (Carlsen et al., 2012). The last type of cross-contamination can occur during sequencing, such as barcode sequencing errors due to low complexity among the index sequences, residual amplicons from previous sequencing runs, or from "index-hopping," a phenomenon when sequencing platforms misassigns indexing reads from one sample to another (Callahan et al., 2016; Eisenhofer et al., 2019; Larsson et al., 2018). Following the "best practices" laid out by Illumina, which includes using dual indexed libraries with unique indexes, storing libraries at -20°C, and removing free adaptors, can reduce sequencing contamination (Illumina Inc., 2017).

The other type of contamination that plagues microbiome work with low biomass samples is exogenous DNA (Eisenhofer et al., 2019). Despite careful laboratory protocols, the introduction of contaminate DNA into samples can still arise from numerous sources, including laboratory reagents, sampling and laboratory environments, plastic consumables, nucleic acid extraction kits, and the researchers themselves (Dahlberg et al., 2019; Eisenhofer et al., 2019; Glassing et al., 2016; Salter et al., 2014; Shen et al., 2006). A list of over 60 common contaminant taxa complied by Eisenhofer and colleagues (2019) from multiple studies analyzing contamination within blank and no-template controls underscores the pervasive problem of contamination in microbiome research. This extensive list is helpful for identifying potential contaminants, but it does little for preventing future contamination as the types and the quantities of contaminants vary across extraction kits and laboratories and can even change over time within the same laboratory (Weyrich et al., 2019). Therefore, to mitigate the negative impact of contamination in low microbial biomass studies, Eisenhofer and colleagues (2019) have presented the RIDE checklist, a set of minimum criteria for investigators, reviewers, and editors, designed to increase the legitimacy of future studies in this rapidly evolving field.

RIDE Checklist (reproduced from Eisenhofer et al., 2019)

- Report the experimental design and approached used to reduce and assess the contributions of contamination
- Include controls to assess contaminant DNA. One of each type of negative control (sampling blanks, DNA extraction blanks, and no-template amplification) must be included per sampling, extraction, or amplification batch
- Determine the level of contamination of comparing biological samples to controls
- Explore contaminant taxa within each study and report their impact on the interpretation of biological samples

Furthermore, the inclusion of a mock community dilution series as an additional control to assess increasing contamination with serially decreasing mock community biomass provides a valuable tool for evaluating filtering parameters in computational analysis (Karstens et al., 2019).

16S rRNA gene as a phylogenetic marker

This widely used approach for characterizing microbial communities employs amplicon sequencing and analysis of a conserved bacterial taxonomic marker—the 16S ribosomal RNA (rRNA) gene (Jo et al., 2016; Meisel et al., 2016b; Miodovnik et al., 2017; Church et al., 2020). The use of the 16S rRNA gene as a phylogenetic marker is inexpensive and efficient (Case et al., 2007; Janda and Abbott, 2007). The gene is present in almost all bacteria (Case et al., 2007; Janda and Abbott, 2007; Yang et al., 2016). Importantly, the 16S rRNA gene contains both highly conserved regions as well as hypervariable regions. The conserved regions allow for the use of "universal primers" to amplify 16S rRNA gene sequences by PCR (Yang et al., 2016). Differences in the sequences of hypervariable regions between conserved portions of the gene allow taxonomic and phylogenetic classification of bacteria. Additionally, the gene's function has largely remained unchanged over time due to a slow rate of evolution (Woese and Fox, 1977; Case et al., 2007; Janda and Abbott, 2007).

However, the 16S rRNA gene amplicon as a phylogenetic marker has drawbacks. This PCR-based method is susceptible to biases stemming from sample preparation, primer affinity, multiple 16S rRNA gene copy variants, and sequencing error rates (Church et al., 2020; Johnson et al., 2019; Starke et al., 2021; Tremblay et al., 2015). Moreover, taxonomic accuracy and resolution is impacted by the length of the sequenced 16S rRNA amplicon. In- silico and sequence-based experiments to assess the accuracy of using the 16S rRNA gene for taxonomic classification show that targeting specific hypervariable regions, i.e., V4, results in less accurate taxonomic resolution compared to full-length sequencing of the gene (Johnson et al., 2019). Until recently, sequencing platforms have been unable to accurately sequence the full length of the gene (Church et al., 2020; Johnson et al., 2019). Third generation sequencing technology (e.g., PacBio and Nanopore) allows for full-length sequencing of the 16S rRNA gene (Johnson et al., 2019). However, the high cost and high rate of sequencing errors limit its practicality (Gwak and Rho, 2020). Consequently, the choice of which of hypervariable regions of the 16S rRNA gene to target, and the ability to sequence the entire 16S rRNA gene, impacts microbial community structure in downstream analyses (Yang et al., 2016; Johnson et al., 2019; Church et al., 2020).

Partial 16S rRNA gene sequencing has low phylogenetic power beyond the genus-level and limited discriminatory power to delineate between bacterial strains (Janda and Abbott, 2007). Historically, 16S rRNA gene amplicon sequences that are >97% similar are clustered together as an operational taxonomic unit (OTU) (Tremblay et al., 2015; Westcott and Schloss, 2015). However, OTU binning poses a taxonomic classification problem for strains that are highly similar, such as *Bacillus globisporus* and *Bacillus psychrophilus*, which share over 99.5% sequence similarity (Janda and Abbott, 2007). Moreover, multiple 16S rRNA gene copies with diverging sequences can exist within a single bacterium, further complicating taxonomic classification (Stackebrandt and Goebel, 1994; Poyart et al., 2001; Johnson et al., 2019; Church et al., 2020). Bacterial genera identified as having multiple 16S gene copies and polymorphisms within these copies include *Bordetella*, *Neisseria*, *Pseudomonas*, *Streptococcus*, *Shigella*, *Bifidobacterium*, *Collinsella*, and *Staphylococcus* (Janda and Abbott, 2007; Johnson et al., 2019; Church et al., 2020).

More recently, the use of amplicon sequence variants (ASVs) derived from 16S rRNA gene amplicon analysis has been used to classify bacteria, as ASVs can differ by just one nucleotide (Porter and Hajibabaei, 2018). Researchers argue ASVs offer a more precise and accurate measurement of sequence variation, and therefore, some microbiome researchers have called for defining taxa based solely on exact nucleotide sequences (Callahan et al., 2017). However, a 2018 study comparing binning approaches (i.e., OTU versus ASV) found that ASV binning provided just a minor improvement in taxonomic resolution over OTU binning, demonstrating that, in general, ribosomal genetic markers can be inadequate molecular targets if high-genetic resolution

of a microbial community is required (Glassman and Martiny, 2018). Nevertheless, OTUs are still widely used in microbiome research, especially for the comparison of diversity across large datasets (Delgado-Baquerizo et al., 2018).

Taxonomic classification can change over time as taxonomic databases and classifiers evolve to reflect newly discovered functional biochemical properties, novel microbiota, and the inclusion of full-length 16S rRNA gene sequences (Gwak and Rho, 2020; Parks et al., 2018). In our work characterizing microbial differences between *Nell2* variants in a knock-out mouse line, we found that the taxonomy of our targeted microbial trait of interest dramatically changed between the time the study began in 2017 and the writing of this thesis. Specifically, the taxonomy of the microbial trait was reclassified from the class *Betaproteobacteria* to the class *Gammaproteobacteria*. Further, there is now disagreement between taxonomic databases regarding the family-level classification of the microbial trait, with one database classifying the bacterium to a family comprised of numerous mammalian commensals, and another classifying the bacterium to a family containing genera frequently labelled as contaminants in microbiome research. This example underscores the need to develop reliable methods and 16S rRNA gene databases for use in microbiome research, while ensuring replicability over time.

Shotgun sequencing offers an alternative method to sequencing the 16S rRNA gene for analyzing microbial communities. This method fragments DNA into small strands for sequencing. Through numerous rounds of DNA fragmentation and sequencing, multiple overlapping sequencing reads (contigs) are produced, which are then programmatically constructed into a single, continuous sequence. Shotgun sequencing is advantageous in that it allows for sequencing the entire breadth of DNA material within a sample (Morgan and Huttenhower, 2012; Reuter et al., 2015; Walsh, 2018). This strategy better enables species and strain-level classification of bacteria compared to partial 16S rRNA gene sequencing, as full-length 16S rRNA gene sequences, including entire bacterial genomes, can be constructed (Walsh, 2018). Moreover, shotgun sequencing data have the potential to resolve functional properties of host-associated microbial communities to better characterize the dynamic interplay between host and microbe (Franzosa et al., 2015; Walsh, 2018).

16S rRNA gene amplicon sequencing is reliable and cost-effective

We performed a systematic comparison of shotgun metagenomic and 16S rRNA partial gene amplicon sequencing for profiling the microbial communities of ten host species (Rausch et al., 2019). The evaluation of 16S rRNA gene sequencing included the sequencing of two gene regions, i.e., V1-V2 versus V3-V4 regions, as well as two PCR amplification procedures, i.e., one-step versus two-step amplification. The ten host species included model organisms, such as *Mus musculus*, as well as *Homo sapiens*. Interestingly, we found that there was not a distinguishable difference in community variation amongst the five methods investigated—16S rRNA gene amplicon sequencing resulted in bacterial profiles that were similar to those derived from shotgun sequencing for all ten host species. For the human samples, we found that the shotgun derived profiles displayed lower alpha diversity compared to both 16S rRNA gene amplicon profiles (V1-V2 compared with V3-V4 sequencing of the hypervariable region), which were similar in terms of bacterial richness and complexity (Rausch et al., 2019).

Trade-offs in costs, biases, functional information, and taxonomic resolution exist between shotgun and 16S rRNA gene amplicon-based sequencing. Shotgun metagenomic sequencing provides more genetic information, allowing for the functional characterization of host- microbe interactions and increased taxonomic resolution (Franzosa et al., 2015; Walsh, 2018). However, its applicability is limited in terms of cost and high-computational demands for analyzing data (Knight et al., 2012; Morgan and Huttenhower, 2012; Walsh, 2018). Long- read sequencing of the entire 16S rRNA gene similarly allows for strain delineation within taxa and provides more

accurate taxonomic classifications (Gwak and Rho, 2020; Johnson et al., 2019). Nevertheless, third generation sequencing technology is still incumbered by cost and sequencing errors (Gwak and Rho, 2020). As such, the profiling of microbial communities with partial 16S rRNA gene amplicon sequencing remains a reliable and cost- effective strategy.

Selected inflammatory diseases herein

Inflammation is a normal physiological process characterized by five cardinal signs: heat (Latin: *calor)*, pain (*dolor*), redness (*rubor*), swelling (*tumor*), and loss of function (*functio laesa*; (Medzhitov, 2010; Takeuchi and Akira, 2010). It is a general biological response of the innate immune system and occurs acutely in response to numerous stimuli and during wound healing. Inflammation becomes chronic when the response is prolonged, resulting in a shift in the type of cells present at the site of inflammation, and when tissue destruction occurs alongside tissue healing. Chronic inflammation occurs in many diseases that contribute to significant morbidity and early mortality (Medzhitov, 2010). This list includes physical diseases, such as autoimmune conditions, asthma, obesity, diabetes, strokes, certain cancers, cardiovascular disease, rheumatoid arthritis, but also psychiatric conditions, including depression, schizophrenia, post-traumatic stress disorder, and eating disorders (Dalton et al., 2018a, 2018b; Slavich, 2015). Inflammation enables survival by removing injurious agents and stimulating healing processes (Takeuchi and Akira, 2010), but aberrant chronic inflammation results in a temporary reduction in tissue function, which can lead disease states (Medzhitov, 2010).

This thesis characterizes resident skin microbiota communities in the context of three diseases associated with chronic inflammation: 1) bullous pemphigoid, an autoimmune blistering disease, 2) anorexia nervosa, an eating disorder whereby affected patients tend to exhibit pro-inflammatory immune responses, and 3) atopic dermatitis, an inflammatory skin disorder that has been shown to be associated with increased expression of the neurodevelopmental gene *Nell2.*

Bullous pemphigoid
Bullous pemphigoid (BP) is the most common autoimmune skin blistering disease in Europe, with an incidence of about 20 new cases per million per year (Joly et al., 2012; Schmidt and Zillikens, 2013; van Beek et al., 2021). The incidence of BP ranges globally from 2.4 to 21.7 new cases per million annually, with the highest reported incidence occurring in the United Kingdom, involving 42.8 new cases per year. BP is considered a disease of the elderly, commonly occurring in patients over the age of 70 years. Accordingly, the incidence of BP is expected to increase in tandem with the aging European population (Schmidt and Zillikens, 2013). It is thus intriguing that a recent multinational study of 9,000 adults showed that skin microbiota are a predictor of advanced age, more so than oral or gut microbiota (Huang et al., 2020).

The hallmark feature of autoimmune blistering diseases such as BP is the disruption of the skin's physical and immune barrier as a consequence of autoantibodies attacking structural proteins (Stevens et al., 2019). Specifically, BP results from autoantibodies targeting hemidesmosomal proteins in the epidermal basement membrane zone, resulting in severe subepidermal blistering that characteristically occurs on the flexor surfaces of the extremities and trunk (Amber et al., 2018; Schmidt and Zillikens, 2013; Stevens et al., 2019). The severity of BP's pruritic skin blisters considerably affects quality of life. Furthermore, the disease is associated with numerous comorbidities and a significantly increased mortality risk (Amber et al., 2018).

Theories for the development of BP include genetic predisposition, various causes of cellular damage, prescription drug reactions, and epitope spreading (Stevens et al., 2019). Missing amongst the theories of induction of autoimmunity in BP is the role of the skin microbiota in host immune defense. An increasing number of studies have revealed that skin microbiota play a crucial role in several aspects of skin biology, including protective immunity through the control of local inflammatory responses, immune cell functioning, and homeostatic immunity (Byrd et al., 2018; Naik et al., 2012; Sanford and Gallo, 2013). Commensal skin microbiota can produce antimicrobial peptides (Byrd et al., 2018). Some microbes have distinct immune system effects. *Staphylococcus epidermidis*, for example, can induce the production of cytokine interleukin-1a

(IL-1a) in the host, which contributes to host defense and skin inflammation. Consequently, disturbances in the homeostasis between the host and skin microbiota may lead to autoimmune blistering diseases afflicting the skin (Ellebrecht et al., 2015; Salava and Lauerma, 2014).

Figure 5. Bullous pemphigoid lesions. Tension bullae typical of bullous pemphigoid disease. Subepidermal blistering tends to occur on flexor surfaces and trunk. The disease is significantly associated with numerous comorbidities and can be fatal.[1]

Anorexia nervosa

Anorexia nervosa (AN) is a psychiatric condition typically affecting females with an estimated lifetime prevalence between 0.5% to 2.0% and a peak age of onset between 13 and 18 years. The hallmark feature of AN is low body weight for age and height due to extreme caloric restriction. AN is complicated by malnutrition that can lead to life-threatening medical consequences as a result of multiple organ failure and immune system dysfunction. AN is the deadliest psychiatric disorder, imparting a significantly increased mortality risk, more so than schizophrenia, bipolar disorder, or unipolar depression (Arcelus et al., 2011). Moreover, amongst adolescents aged 15 to 24 years, the mortality risk from anorexia nervosa is greater than for any other serious health disease, including asthma or type I diabetes
(Schmidt et al., 2016).

Nutrition is closely tied to immune system functioning (Calder and Jackson, 2000). Starvation, malnutrition, altered dietary patterns, and single-nutrient deficiencies are all important causes of impaired immune functioning that can lead to chronic inflammation and recurrent infections (Bourke et al., 2016; Calder and Jackson, 2000). Children suffering from malnutrition chiefly die from “common infections,” suggesting a link between nutrition status and immune functioning (Bourke et al., 2016; Kim et al., 2014; Rytter et al., 2014). In the context of childhood malnutrition typically seen in developing nations, but also in clinical malnutrition more generally, Bourke and colleagues (2016) have proposed immune system dysfunction as both a driver and repercussion of undernutrition.

In line with this proposed hypothesis, numerous studies have reported altered immunity profiles in patients suffering from anorexia nervosa (Dalton et al., 2018a; Omodei et al., 2015; Słotwiński and Słotwiński, 2017). Moreover, AN is associated with dermatological changes, including xerosis (dry skin), increased acne, slower wound healing, generalized pruritis, and seborrheic dermatitis, all of which are associated with an increased risk of infection if the skin barrier is compromised (Augustin et al., 2019; Guo and DiPietro, 2010; Strumìa et al., 2001; Westmoreland et al., 2016). Yet, clinical observations report an absence of infections in patients with anorexia nervosa as well as delayed or reduced physiological responses to infection, in general (Bowers and Eckert, 1978; Brown et al., 2008, 2005; Nova and Marcos, 2006). Paradoxically, cell-

mediated immunity appears to be amplified in AN, with several studies reporting an increase in T-cell proliferation and inflammatory cytokine production, including interleukin-1 (IL-1), interleukin-6 (IL-6), and tumor-necrosis factor (TNF), when compared to healthy controls or to innate immunity responses in primary malnutrition, where immune function is suppressed (Dalton et al., 2018a, 2018b; Gibson and Mehler, 2019; Słotwiński and Słotwiński, 2017). An earlier investigation conducted by Omodei and colleagues (2015) found that immune cell abundances are reduced in anorexia nervosa, but that these patients also exhibited a greater antioxidant, stress resistance, and anti- inflammatory profiles compared to controls. While the link between anorexia nervosa and altered immune function has yet to be fully elucidated, an altered skin microbial profile in patients with AN might contribute to augmented skin immunity, thereby increasing resistance to skin infections.

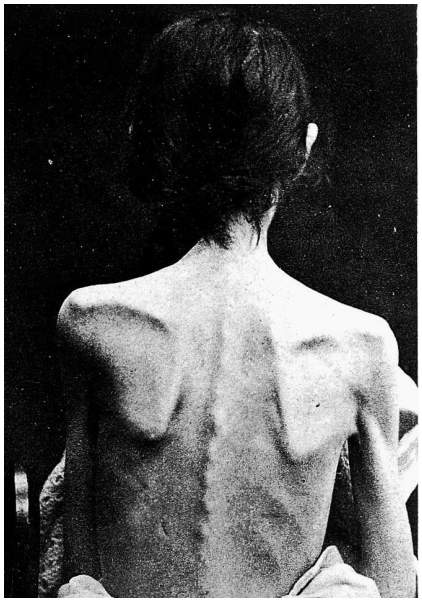


Figure 6. A patient with anorexia nervosa (AN). Muscle wasting and extreme starvation seen in AN. Malnutrition resulting from AN is life-threatening—it is considered the deadliest psychiatric disorder. AN is associated with changes in the skin, including increased dryness, increased acne, slow wound healing, generalized pruritis, and seborrheic dermatitis. These conditions are associated with an increased risk of infection if the skin barrier is compromised.[2]

Atopic dermatitis

Atopic dermatitis (AD) is the most common chronic inflammatory skin disease in developed nations (Egawa and Kabashima, 2016; Kim et al., 2019). A recent systematic review of the prevalence and incidence of atopic dermatitis found that among children, the prevalence of AD is about 20% and between 7% and 14%, among adults depending on the country, with developed nations representing the higher end of the spectrum. Generally, the onset of AD occurs in childhood, with remission commonly occurring in adolescence (Bylund et al., 2020). Persistent or severe AD significantly affects quality of life due to symptoms of intense itching (pruritis), inflammation, and chronic skin infections (Na et al., 2019).

The pathophysiology of AD is multifactorial and not completely understood (Egawa and Kabashima, 2016; Kim et al., 2019). The scientific community generally agrees that environmental conditions cause AD for those with genetic susceptibility (Løset et al., 2019). Several studies have found that genetic risk factors are connected to the development of AD. The most notable of these is the 2006 discovery of loss-of-function mutations of the *FLG* gene

encoding the filaggrin protein, located on chromosome 1q21.3 in the epidermal differentiation complex (Løset et al., 2019). This finding substantially changed the landscape of AD pathogenesis research, which thereafter shifted away from solely focusing on immunological features, such as T cell imbalance, to also concentrating on epidermal barrier dysfunction, which can be sparked by genetic mutations (Elias and Steinhoff, 2008; Løset et al., 2019).
Evolving evidence demonstrates that skin barrier dysfunction plays a critical role in the emergence of AD and that numerous factors contribute to the breakdown of the skin barrier, including immune dysregulation, a deficiency of antimicrobial peptides, as well as defects in epidermal barrier proteins such as filaggrin (Egawa and Kabashima, 2016; Kim et al., 2019). An alteration in resident skin microbiota might also contribute to skin barrier dysfunction via altered immune functions, inflammation, and infection, and moreover, biological host factors may spur further microbial changes. (Chng et al., 2016; Clausen et al., 2018; Egawa and Kabashima, 2016; Kim et al., 2019; Nakatsuji et al., 2017; Salava and Lauerma, 2014).

The skin of AD patients shows decreased expression of AMPs, which has been linked to increased susceptibility to colonization by potentially pathogenic bacteria, such as *Staphylococcus aureus* (Nakatsuji et al., 2017). Colonization by *S. aureus* can exacerbate the positive feedback loop driving disease progression via immune cell dysfunction, further reduction in AMP expression, the development of allergic co-morbidities, and increased destruction of the skin barrier (Nakatsuji et al., 2017). Moreover, mice experimentally colonized by *S. aureus* will develop AD-like lesions, likely caused by the exploitation of barrier defects and by triggering inflammation (Kobayashi et al., 2015; Nakatsuji et al., 2017).

Figure 7. Atopic dermatitis (AD) lesion on the arm. A typical AD lesion with associated skin inflammation and signs of pruritus. Skin barrier dysfunction plays a critical role in AD pathophysiology. Numerous factors contribute to the breakdown of the skin barrier, including skin immunity dysregulation, antimicrobial peptide deficiency, and genetic defects affecting key epidermal barrier proteins like filaggrin.[3]

An unlikely genetic candidate that perhaps links skin microbiota and skin barrier dysfunction is the *Nell2* (neural epidermal growth factor-like 2) gene, which encodes the enzyme kinase C-binding protein NELL2 in humans (Watanabe et al., 1996). *Nell2* is predominately expressed in neural tissues and has been reported to play a central role in the proliferation and differentiation of neural cells (Kim et al., 2014). *Nell2* is also differentially expressed in the epidermis of patients suffering from atopic dermatitis (Kamsteeg et al., 2010). Interestingly, a 2011 study utilizing a human skin equivalent exhibiting atopic dermatitis characteristics demonstrated an increase in *Nell2* expression after stimulation with Th2 cytokines, which are associated with atopy, but not with the stimulation of Th17-related or psoriasis-related cytokines (Berger, 2000; Kamsteeg et al.,

2011). More recently, using the 15th generation of an advanced intercross line and a QTL mapping approach, Belheouane and colleagues (2017) identified numerous genomic loci associated with skin microbial abundances, with some containing single immune- and/or skin cancer-related genes. One exceptional candidate region was associated with unclassified *Betaproteobacteria* and contained only one gene: *Nell2*.

While the relationship between *Nell2* and AD is only recently characterized, it is intriguing that Urashima and colleagues (1998) reported an increased number of cutaneous free nerve endings in the epidermis of patients with AD, perhaps contributing to the severe pruritis that epitomizes atopic dermatitis (Urashima and Mihara, 1998). It is therefore plausible that increased *Nell2* expression in AD mediates cutaneous nerve fiber proliferation in the epidermis, and furthermore, that there is a host-microbial relationship between *Nell2* expression and the bacteria that produce proteases that cause or exacerbate epidermal barrier dysfunction via inflammation, infection, and pruritis in the context of inflammatory skin disease (Koziel and Potempa, 2013; Steinhoff et al., 2003; Sung Kim and Yosipovitch, 2020).

References

Acinas SG, Sarma-Rupavtarm R, Klepac-Ceraj V, Polz MF. PCR-induced sequence artifacts and bias: insights from comparison of two 16S rRNA clone libraries constructed from the same sample. Appl Environ Microbiol. 2005;71:8966–9. https://doi.org/10.1128/AEM.71.12.8966-8969.2005.

Amber KT, Murrell DF, Schmidt E, Joly P, Borradori L. Autoimmune Subepidermal Bullous Diseases of the Skin and Mucosae: Clinical Features, Diagnosis, and Management. Clin Rev Allergy Immunol. 2018;54:26–51. https://doi.org/10.1007/s12016-017-8633-4.

Arcelus J, Mitchell AJ, Wales J, Nielsen S. Mortality Rates in Patients with Anorexia Nervosa and Other Eating Disorders: A Meta-analysis of 36 Studies. Archives of General Psychiatry. 2011;68:724–31. https://doi.org/10.1001/archgenpsychiatry.2011.74.

Augustin M, Wilsmann-Theis D, Körber A, Kerscher M, Itschert G, Dippel M, et al. Diagnosis and treatment of xerosis cutis – a position paper. JDDG: Journal Der Deutschen Dermatologischen Gesellschaft. 2019;17:3–33. https://doi.org/10.1111/ddg.13906.

Awany D, Allali I, Dalvie S, Hemmings S, Mwaikono KS, Thomford NE, et al. Host and Microbiome Genome- Wide Association Studies: Current State and Challenges. Frontiers in Genetics. 2019;9. https://doi.org/10.3389/fgene.2018.00637

van Beek N, Weidinger A, Schneider SW, Kleinheinz A, Gläser R, Holtsche MM, et al. Incidence of pemphigoid diseases in Northern Germany in 2016 – first data from the Schleswig-Holstein Registry of Autoimmune Bullous Diseases. J Eur Acad Dermatol Venereol. 2021. https://doi.org/10.1111/jdv.17107.

Belheouane M, Gupta Y, Künzel S, Ibrahim S, Baines JF. Improved detection of gene-microbe interactions in the mouse skin microbiota using high-resolution QTL mapping of 16S rRNA transcripts. Microbiome. 2017;5:59. https://doi.org/10.1186/s40168-017-0275-5.

Belheouane M, Hermes BM, Van Beek N, Benoit S, Bernard P, Drenovska K, et al. Characterization of the skin microbiota in bullous pemphigoid patients and controls reveals novel microbial indicators of disease. J Adv Res 2022. https://doi.org/10.1016/j.jare.2022.03.019.

Belkaid Y, Segre JA. Dialogue between skin microbiota and immunity. Science 2014;346:954–9. https://doi.org/10.1126/science.1260144.

Berger A. Th1 and Th2 responses: what are they? BMJ. 2000;321:424. https://doi.org/10.1136/bmj.321.7258.424

Blaxter M, Mann J, Chapman T, Thomas F, Whitton C, Floyd R, et al. Defining operational taxonomic units using DNA barcode data. Philos Trans R Soc Lond B Biol Sci. 2005;360:1935–43. https://doi.org/10.1098/rstb.2005.1725.

Blekhman R, Goodrich JK, Huang K, Sun Q, Bukowski R, Bell JT, et al. Host genetic variation impacts microbiome composition across human body sites. Genome Biol. 2015;16. https://doi.org/10.1186/s13059-015- 0759-1.

Bosch TCG, McFall-Ngai MJ. Metaorganisms as the new frontier. Zoology. 2011;114:185–90.

https://doi.org/10.1016/j.zool.2011.04.001.

Bourke CD, Berkley JA, Prendergast AJ. Immune Dysfunction as a Cause and Consequence of Malnutrition. Trends Immunol. 2016;37:386–98. https://doi.org/10.1016/j.it.2016.04.003.

Bowers TK, Eckert E. Leukopenia in anorexia nervosa. Lack of increased risk of infection. Arch Intern Med. 1978;138:1520–3. https://doi:10.1001/archinte.1978.03630350050015

Brown RF, Bartrop R, Beumont P, Birmingham CL. Bacterial infections in anorexia nervosa: delayed recognition increases complications. Int J Eat Disord. 2005;37:261–5. https://doi.org/10.1002/eat.20135.

Brown RF, Bartrop R, Birmingham CL. Immunological disturbance and infectious disease in anorexia nervosa: a review. Acta Neuropsychiatr. 2008;20:117–28. https://doi.org/10.1111/j.1601-5215.2008.00286.x.

Bylund S, Kobyletzki LB, Svalstedt M, Svensson Å. Prevalence and Incidence of Atopic Dermatitis: A Systematic Review. Acta Derm Venereol. 2020;100:adv00160. https://doi.org/10.2340/00015555-3510.

Byrd AL, Belkaid Y, Segre JA. The human skin microbiome. Nat Rev Microbiols. 2018;16:143–55. https://doi.org/10.1038/nrmicro.2017.157.

Calder PC, Jackson AA. Undernutrition, infection and immune function. Nutr Res Rev. 2000;13:3–29. https://doi.org/10.1079/095442200108728981.

Callahan BJ, McMurdie PJ, Holmes SP. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. ISME J. 2017;11:2639–43. https://doi.org/10.1038/ismej.2017.119.

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13:581–3. https://doi.org/10.1038/nmeth.3869.

Carlsen T, Aas AB, Lindner D, Vrålstad T, Schumacher T, Kauserud H. Don't make a mista(g)ke: is tag switching an overlooked source of error in amplicon pyrosequencing studies? Fungal Ecology. 2012;5:747–9. https://doi.org/10.1016/j.funeco.2012.06.003.

Caruso V, Song X, Asquith M, Karstens L. Performance of Microbiome Sequence Inference Methods in Environments with Varying Biomass. Msystems. 2019;4:e00163-18. https://doi.org/10.1128/mSystems.00163- 18.

Case RJ, Boucher Y, Dahllöf I, Holmström C, Doolittle WF, Kjelleberg S. Use of 16S rRNA and rpoB Genes as Molecular Markers for Microbial Ecology Studies. Appl Environ Microbiol. 2007;73:278–88. https://doi.org/10.1128/AEM.01177-06.

Chng KR, Tay ASL, Li C, Ng AHQ, Wang J, Suri BK, et al. Whole metagenome profiling reveals skin microbiome-dependent susceptibility to atopic dermatitis flare. Nat Microbiol. 2016;1:1–10. https://doi.org/10.1038/nmicrobiol.2016.106.

Church DL, Cerutti L, Gürtler A, Griener T, Zelazny A, Emler S. Performance and Application of 16S rRNA Gene Cycle Sequencing for Routine Identification of Bacteria in the Clinical

Microbiology Laboratory. Clin Microbiol Rev. 2020;33:e00053-19. https://doi.org/10.1128/CMR.00053-19.

Clausen M-L, Agner T, Lilje B, Edslev SM, Johannesen TB, Andersen PS. Association of Disease Severity with Skin Microbiome and Filaggrin Gene Mutations in Adult Atopic Dermatitis. JAMA Dermatol. 2018;154:293–300. https://doi.org/10.1001/jamadermatol.2017.5440.

Cogen AL, Nizet V, Gallo RL. Skin microbiota: a source of disease or defence? Br J Dermatol. 2008;158:442–55. https://doi.org/10.1111/j.1365-2133.2008.08437.x.

Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. Bacterial Community Variation in Human Body Habitats Across Space and Time. Science. 2009;326:1694–7. https://doi.org/10.1126/science.1177486.

Dahlberg J, Sun L, Persson Waller K, Östensson K, McGuire M, Agenäs S, et al. Microbiota data from low biomass milk samples is markedly affected by laboratory and reagent contamination. PLoS One. 2019;14:e0218257. https://doi.org/10.1371/journal.pone.0218257.

Dalton B, Bartholdy S, Robinson L, Solmi M, Ibrahim MAA, Breen G, et al. A meta-analysis of cytokine concentrations in eating disorders. J Psychiatr Res. 2018a;103:252–64. https://doi.org/10.1016/j.jpsychires.2018.06.002.

Dalton B, Campbell IC, Chung R, Breen G, Schmidt U, Himmerich H. Inflammatory Markers in Anorexia Nervosa: An Exploratory Study. Nutrients. 2018b;10:1573. https://doi.org/10.3390/nu10111573.

De Filippo C, Cavalieri D, Di Paola M, Ramazzotti M, Poullet JB, Massart S, et al. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. Proc Natl Acad Sci USA. 2010;107:14691–6. https://doi.org/10.1073/pnas.1005963107.

Delgado-Baquerizo M, Oliverio AM, Brewer TE, Benavent-González A, Eldridge DJ, Bardgett RD, et al. A global atlas of the dominant bacteria found in soil. Science. 2018;359:320–5. https://doi.org/10.1126/science.aap9516.

Egawa G, Kabashima K. Multifactorial skin barrier deficiency and atopic dermatitis: Essential topics to prevent the atopic march. J Allergy Clin Immunol. 2016;138:350-358.e1. https://doi.org/10.1016/j.jaci.2016.06.002.

Eisenhofer R, Minich JJ, Marotz C, Cooper A, Knight R, Weyrich LS. Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. Trends in Microbiol. 2019;27:105–17. https://doi.org/10.1016/j.tim.2018.11.003.

Elias PM, Steinhoff M. “Outside-to-Inside” (and Now Back to “Outside”) Pathogenic Mechanisms in Atopic Dermatitis. J Invest Dermatol 2008;128:1067–70. https://doi.org/10.1038/jid.2008.88.

Ellebrecht CT, Srinivas G, Bieber K, Banczyk D, Kalies K, Künzel S, et al. Skin microbiota-associated inflammation precedes autoantibody induced tissue damage in experimental epidermolysis bullosa acquisita. J Autoimmun. 2015. https://doi.org/10.1016/j.jaut.2015.08.007.

Fierer N, Lauber CL, Zhou N, McDonald D, Costello EK, Knight R. Forensic identification using

skin bacterial communities. PNAS. 2010;107:6477–81. https://doi.org/10.1073/pnas.1000162107.

Franzosa EA, Hsu T, Sirota-Madi A, Shafquat A, Abu-Ali G, Morgan XC, et al. Sequencing and beyond: integrating molecular 'omics for microbial community profiling. Nat Rev Microbiol. 2015;13:360–72. https://doi.org/10.1038/nrmicro3451.

Fuchs E, Raghavan S. Getting under the skin of epidermal morphogenesis. Nat Rev Genet. 2002;3:199–209. https://doi.org/10.1038/nrg758.

Gensollen T, Iyer SS, Kasper DL, Blumberg RS. How colonization by microbiota in early life shapes the immune system. Science. 2016;352:539–44. https://doi.org/10.1126/science.aad9378.

Gibson D, Mehler PS. Anorexia Nervosa and the Immune System—A Narrative Review. J Clin Med. 2019;8. https://doi.org/10.3390/jcm8111915.

Glassing A, Dowd SE, Galandiuk S, Davis B, Chiodini RJ. Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. Gut Pathog. 2016;8:24. https://doi.org/10.1186/s13099-016-0103-7.

Glassman SI, Martiny JBH. Broadscale Ecological Patterns Are Robust to Use of Exact Sequence Variants versus Operational Taxonomic Units. MSphere. 2018;3:e00148-18. https://doi.org/10.1128/mSphere.00148-18.

Goodrich JK, Davenport ER, Beaumont M, Jackson MA, Knight R, Ober C, et al. Genetic determinants of the gut microbiome in UK Twins. Cell Host Microbe. 2016;19:731–43. https://doi.org/10.1016/j.chom.2016.04.017.

Goodrich JK, Waters JL, Poole AC, Sutter JL, Koren O, Blekhman R, et al. Human Genetics Shape the Gut Microbiome. Cell. 2014;159:789–99. https://doi.org/10.1016/j.cell.2014.09.053.

Graham P, X. Lin S, L. Larson E. A U.S. Population-Based Survey of Staphylococcus aureus Colonization. Ann Intern Med. 2006;144(5):318-25. doi: 10.7326/0003-4819-144-5-200603070-00006.

Grice EA, Kong HH, Conlan S, Deming CB, Davis J, Young AC, et al. Topographical and Temporal Diversity of the Human Skin Microbiome. Science. 2009;324:1190–2. https://doi.org/10.1126/science.1171700.

Grice EA, Kong HH, Renaud G, Young AC, Bouffard GG, Blakesley RW, et al. A diversity profile of the human skin microbiota. Genome Res. 2008;18:1043–50. https://doi.org/10.1101/gr.075549.107.

Grice EA, Segre JA. The skin microbiome. Nat Rev Microbiol. 2011;9:244–53. https://doi.org/10.1038/nrmicro2537.

Guo S, DiPietro LA. Factors Affecting Wound Healing. J Dent Res. 2010;89:219–29. https://doi.org/10.1177/0022034509359125.

Gupta VK, Paul S, Dutta C. Geography, Ethnicity or Subsistence-Specific Variations in Human Microbiome Composition and Diversity. Front Microbiol. 2017;8. https://doi.org/10.3389/fmicb.2017.01162.

Gwak H-J, Rho M. Data-Driven Modeling for Species-Level Taxonomic Assignment From 16S rRNA: Application to Human Microbiomes. Frontiers in Microbiology. 2020;11. https://doi.org/10.3389/fmicb.2020.570825.

Huang S, Haiminen N, Carrieri A-P, Hu R, Jiang L, Parida L, et al. Human Skin, Oral, and Gut Microbiomes Predict Chronological Age. MSystems 2020;5:e00630-19, /msystems/5/1/msys.00630-19.atom. https://doi.org/10.1128/mSystems.00630-19.

Iacob S, Iacob DG, Luminos LM. Intestinal Microbiota as a Host Defense Mechanism to Infectious Threats. Front Microbiol. 2018;9:3328. https://doi.org/10.3389/fmicb.2018.03328.

Illumina Inc. Effects of Index Misassignment on Multiplexing and Downstream Analysis. Publication No 770- 2017-004-C QB #5420. 2017:4.

Jacobs MC, Haak BW, Hugenholtz F, Wiersinga WJ. Gut microbiota and host defense in critical illness. Curr Opin Crit Care. 2017;23:257–63. https://doi.org/10.1097/MCC.0000000000000424.

Janda JM, Abbott SL. 16S rRNA Gene Sequencing for Bacterial Identification in the Diagnostic Laboratory: Pluses, Perils, and Pitfalls. J Clin Microbiol. 2007;45:2761–4. https://doi.org/10.1128/JCM.01228-07.

Jo J-H, Kennedy EA, Kong HH. Bacterial 16S ribosomal RNA gene sequencing in cutaneous research. J Invest Dermatol. 2016;136:e23–7. https://doi.org/10.1016/j.jid.2016.01.005.

Johnson JS, Spakowicz DJ, Hong B-Y, Petersen LM, Demkowicz P, Chen L, et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. Nat Commun. 2019;10:5029. https://doi.org/10.1038/s41467-019-13036-1.

Joly P, Baricault S, Sparsa A, Bernard P, Bédane C, Duvert-Lehembre S, et al. Incidence and Mortality of Bullous Pemphigoid in France. J Invest Dermatol. 2012;132:1998–2004. https://doi.org/10.1038/jid.2012.35.

Kamsteeg M, Bergers M, de Boer R, Zeeuwen PLJM, Hato SV, Schalkwijk J, et al. Type 2 Helper T-Cell Cytokines Induce Morphologic and Molecular Characteristics of Atopic Dermatitis in Human Skin Equivalent. Am J Pathol. 2011;178:2091–9. https://doi.org/10.1016/j.ajpath.2011.01.037.

Kamsteeg M, Jansen P a. M, Vlijmen-Willems IMJJV, Erp PEJV, Rodijk-Olthuis D, Valk PGVD, et al. Molecular diagnostics of psoriasis, atopic dermatitis, allergic contact dermatitis and irritant contact dermatitis. Br J Dermatol. 2010;162. https://doi.org/10.1111/j.1365-2133.2009.09547.x.

Karstens L, Asquith M, Davin S, Fair D, Gregory WT, Wolfe AJ, et al. Controlling for Contaminants in Low- Biomass 16S rRNA Gene Sequencing Experiments. MSystems .2019;4:e00290-19. https://doi.org/10.1128/mSystems.00290-19.

Kim, D, Hofstaedter, CE, Zhao, C, et al. Optimizing methods and dodging pitfalls in microbiome research. Microbiome. 2017;5:52. https://doi.org/10.1186/s40168-017-0267-5

Kim DH, Kim HR, Choi EJ, Kim DY, Kim KK, Kim BS, et al. Neural Epidermal Growth Factor-Like Like Protein 2 (NELL2) Promotes Aggregation of Embryonic Carcinoma P19 Cells by Inducing N-Cadherin Expression. PLoS One. 2014;9:e85898.

https://doi.org/10.1371/journal.pone.0085898.

Kim J, Kim BE, Leung DYM. Pathophysiology of atopic dermatitis: Clinical implications. Allergy Asthma Proc. 2019;40:84–92. https://doi.org/10.2500/aap.2019.40.4202.

Klasić M, Markulin D, Vojta A, Samaržija I, Biruš I, Dobrinić P, et al. Promoter methylation of the MGAT3 and BACH2 genes correlates with the composition of the immunoglobulin G glycome in inflammatory bowel disease. Clin Epigenetics. 2018;10:75. https://doi.org/10.1186/s13148-018-0507-y.

Klein RJ. Power analysis for genome-wide association studies. BMC Genet .2007;8:58. https://doi.org/10.1186/1471-2156-8-58.

Knight R, Jansson J, Field D, Fierer N, Desai N, Fuhrman JA, et al. Unlocking the potential of metagenomics through replicated experimental design. Nat Biotechnol. 2012;30:513–20. https://doi.org/10.1038/nbt.2235.

Kobayashi T, Glatz M, Horiuchi K, Kawasaki H, Akiyama H, Kaplan DH, et al. Dysbiosis and Staphylococcus aureus colonization drives inflammation in atopic dermatitis. Immunity. 2015;42:756–66. https://doi.org/10.1016/j.immuni.2015.03.014.

Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, et al. Succession of microbial consortia in the developing infant gut microbiome. Proc Natl Acad Sci USA. 2011;108:4578–85. https://doi.org/10.1073/pnas.1000081107.

Kong HH. Details Matter: Designing Skin Microbiome Studies. J Invest Dermatol. 2016;136:900–2. https://doi.org/10.1016/j.jid.2016.03.004.

Kong HH, Andersson B, Clavel T, Common JE, Jackson SA, Olson ND, et al. Performing Skin Microbiome Research: A Method to the Madness. J Invest Dermatol. 2017;137:561–8. https://doi.org/10.1016/j.jid.2016.10.033.

Kopylova E, Navas-Molina JA, Mercier C, Xu ZZ, Mahé F, He Y, et al. Open-Source Sequence Clustering Methods Improve the State of the Art. Msystems. 2016. https://doi.org/10.1128/mSystems.00003-15.

Koziel J, Potempa J. Protease-armed bacteria in the skin. Cell Tissue Res. 2013;351:325–37. https://doi.org/10.1007/s00441-012-1355-2.

Kunin V, Engelbrektson A, Ochman H, Hugenholtz P. Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. Environ Microbiol 2010;12:118–23. https://doi.org/10.1111/j.1462-2920.2009.02051.x.

de Lange K, Moutsianas L, Lee J, Lamb C, Luo Y, Na K, et al. Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. Nat Genet. 2017;49. https://doi.org/10.1038/ng.3760.

Larsson AJM, Stanley G, Sinha R, Weissman IL, Sandberg R. Computational correction of index switching in multiplexed sequencing libraries. Nat Methods. 2018;15:305–7. https://doi.org/10.1038/nmeth.4666.

Lauber CL, Zhou N, Gordon JI, Knight R, Fierer N. Effect of storage conditions on the

assessment of bacterial community structure in soil and human-associated samples. FEMS Microbiol Lett. 2010;307:80–6. https://doi.org/10.1111/j.1574-6968.2010.01965.x.

Lederberg J, McCray AT. 'Ome Sweet 'Omics-- A Genealogical Treasury of Words. The Scientist. 2001;15:8.

Ley RE, Peterson DA, Gordon JI. Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell. 2006;124:837–48. https://doi.org/10.1016/j.cell.2006.02.017.

Liu JZ, van Sommeren S, Huang H, Ng SC, Alberts R, Takahashi A, et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. Nat Genet. 2015;47:979–86. https://doi.org/10.1038/ng.3359.

Løset M, Brown SJ, Saunes M, Hveem K. Genetics of Atopic Dermatitis: From DNA Sequence to Clinical Relevance. Dermatology. 2019;235:355–64. https://doi.org/10.1159/000500402.

Marchesi JR, Ravel J. The vocabulary of microbiome research: a proposal. Microbiome. 2015;3:31. https://doi.org/10.1186/s40168-015-0094-5.

Masiuk H, Wcisłek A, Jursa-Kulesza J. Determination of nasal carriage and skin colonization, antimicrobial susceptibility and genetic relatedness of Staphylococcus aureus isolated from patients with atopic dermatitis in Szczecin, Poland. BMC Infect Dis. 2021;21:701. https://doi.org/10.1186/s12879-021-06382-3.

McGovern DPB, Jones MR, Taylor KD, Marciante K, Yan X, Dubinsky M, et al. Fucosyltransferase 2 (FUT2) non-secretor status is associated with Crohn's disease. Hum Mol Genet. 2010;19:3468–76. https://doi.org/10.1093/hmg/ddq248.

Medzhitov R. Inflammation 2010: new adventures of an old flame. Cell. 2010;140:771–6. https://doi.org/10.1016/j.cell.2010.03.006.

Meisel JS, Hannigan GD, Tyldsley AS, SanMiguel AJ, Hodkinson BP, Zheng Q, et al. Skin Microbiome Surveys Are Strongly Influenced by Experimental Design. J Invest Dermatol. 2016a;136:947–56. https://doi.org/10.1016/j.jid.2016.01.016.

Meisel JS, Hannigan GD, Tyldsley AS, SanMiguel AJ, Hodkinson BP, Zheng Q, et al. Skin Microbiome Surveys Are Strongly Influenced by Experimental Design. J Invest Dermatol. 2016b;136:947–56. https://doi.org/10.1016/j.jid.2016.01.016.

Miodovnik M, Künstner A, Langan EA, Zillikens D, Gläser R, Sprecher E, et al. A distinct cutaneous microbiota profile in autoimmune bullous disease patients. Exp Dermatol. 2017. https://doi.org/10.1111/exd.13357.

Molteni M, Gemma S, Rossetti C. The Role of Toll-Like Receptor 4 in Infectious and Noninfectious Inflammation. Mediators Inflamm. 2016;2016:6978936. https://doi.org/10.1155/2016/6978936.

Momozawa Y, Dmitrieva J. IBD risk loci are enriched in multigenic regulatory modules encompassing putative causative genes. Nat Commun. 2018;9. https://doi.org/DOI: 10.1038/s41467-018-04365-8.

Morgan XC, Huttenhower C. Chapter 12: Human Microbiome Analysis. PLoS Comput Biol.

2012;8:e1002808. https://doi.org/10.1371/journal.pcbi.1002808.

Na CH, Chung J, Simpson EL. Quality of Life and Disease Impact of Atopic Dermatitis and Psoriasis on Children and Their Families. Children (Basel). 2019;6:133. https://doi.org/10.3390/children6120133.

Naik S, Bouladoux N, Wilhelm C, Molloy MJ, Salcedo R, Kastenmuller W, et al. Compartmentalized Control of Skin Immunity by Resident Commensals. Science. 2012;337:1115–9. https://doi.org/10.1126/science.1225152.

Nakatsuji T, Chen TH, Narala S, Chun KA, Two AM, Yun T, et al. Antimicrobials from human skin commensal bacteria protect against Staphylococcus aureus and are deficient in atopic dermatitis. Sci Transl Med. 2017;9. https://doi.org/10.1126/scitranslmed.aah4680.

Nakatsuji T, Chiang H-I, Jiang SB, Nagarajan H, Zengler K, Gallo RL. The microbiome extends to subepidermal compartments of normal skin. Nat Commun. 2013;4:1431. https://doi.org/10.1038/ncomms2441.

Nomura I, Gao B, Boguniewicz M, Darst MA, Travers JB, Leung DYM. Distinct patterns of gene expression in the skin lesions of atopic dermatitis and psoriasis. J Allergy Clin Immunol. 2003;112:1195–202. https://doi.org/10.1016/j.jaci.2003.08.049.

Nova E, Marcos A. Immunocompetence to assess nutritional status in eating disorders. Expert Rev Clin Immunol. 2006;2:433–44. https://doi.org/10.1586/1744666X.2.3.433.

Oh J, Byrd AL, Deming C, Conlan S, Kong HH, Segre JA. Biogeography and individuality shape function in the human skin metagenome. Nature. 2014;514:59–64. https://doi.org/10.1038/nature13786.

Oh J, Byrd AL, Park M, Kong HH, Segre JA. Temporal Stability of the Human Skin Microbiome. Cell. 2016;165:854–66. https://doi.org/10.1016/j.cell.2016.04.008.

Omodei D, Pucino V, Labruna G, Procaccini C, Galgani M, Perna F, et al. Immune-metabolic profiling of anorexic patients reveals an anti-oxidant and anti-inflammatory phenotype. Metabolism. 2015;64:396–405. https://doi.org/10.1016/j.metabol.2014.10.025.

Org E, Org E, Parks BW, Joo JWJ, Emert B, Schwartzman W, et al. Genetic and environmental control of host- gut microbiota interactions. Genome Res. 2015;25:1558–69. https://doi.org/10.1101/gr.194118.115.

Ormerod KL, Wood DLA, Lachner N, Gellatly SL, Daly JN, Parsons JD, et al. Genomic characterization of the uncultured Bacteroidales family S24-7 inhabiting the guts of homeothermic animals. Microbiome. 2016;4:36. https://doi.org/10.1186/s40168-016-0181-2.

Paik Y-H, Schwabe RF, Bataller R, Russo MP, Jobin C, Brenner DA. Toll-Like receptor 4 mediates inflammatory signaling by bacterial lipopolysaccharide in human hepatic stellate cells. Hepatology. 2003;37:1043–55. https://doi.org/10.1053/jhep.2003.50182.

Palmer C, Bik EM, DiGiulio DB, Relman DA, Brown PO. Development of the Human Infant Intestinal Microbiota. PLoS Biol 2007;5:e177. https://doi.org/10.1371/journal.pbio.0050177.

Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, et al. A

standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol. 2018;36:996–1004. https://doi.org/10.1038/nbt.4229.

Peacock SJ, de Silva I, Lowy FD. What determines nasal carriage of Staphylococcus aureus? Trends Microbiol. 2001;9:605–10. https://doi.org/10.1016/s0966-842x(01)02254-5.

Porter TM, Hajibabaei M. Scaling up: A guide to high-throughput genomic approaches for biodiversity analysis. Mol Ecol. 2018;27:313–38. https://doi.org/10.1111/mec.14478.

Poyart C, Quesne G, Boumaila C, Trieu-Cuot P. Rapid and Accurate Species-Level Identification of Coagulase- Negative Staphylococci by Using the sodA Gene as a Target. J Clin Microbiol .2001;39:4296–301. https://doi.org/10.1128/JCM.39.12.4296-4301.2001.

Proksch E, Brandner JM, Jensen J-M. The skin: an indispensable barrier. Exp Dermatol. 2008;17:1063–72. https://doi.org/10.1111/j.1600-0625.2008.00786.x.

Rausch P, Rühlemann M, Hermes BM, Doms S, Dagan T, Dierking K, et al. Comparative analysis of amplicon and metagenomic sequencing methods reveals key features in the evolution of animal metaorganisms. Microbiome. 2019;7:133. https://doi.org/10.1186/s40168-019-0743-1.

Reuter JA, Spacek DV, Snyder MP. High-Throughput Sequencing Technologies. Mol Cell. 2015;58:586–97. https://doi.org/10.1016/j.molcel.2015.05.004.

Røysamb E, Tambs K. The beauty, logic and limitations of twin studies. Norsk Epidemiologi. 2016;26. https://doi.org/10.5324/nje.v26i1-2.2014.

Rühlemann MC, Hermes BM, Bang C, Doms S, Moitinho-Silva L, Thingholm LB, et al. Genome-wide association study in 8,956 German individuals identifies influence of ABO histo-blood groups on gut microbiome. Nat Genet. 2021;53:147–55. https://doi.org/10.1038/s41588-020-00747-1.

Rytter MJH, Kolte L, Briend A, Friis H, Christensen VB. The Immune System in Children with Malnutrition— A Systematic Review. PLoS One. 2014;9:e105017. https://doi.org/10.1371/journal.pone.0105017.

Salava A, Lauerma A. Role of the skin microbiome in atopic dermatitis. Clin Transl Allergy. 2014;4:33. https://doi.org/10.1186/2045-7022-4-33.

Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biology. 2014;12:87. https://doi.org/10.1186/s12915-014-0087-z.

Sanford JA, Gallo RL. Functions of the skin microbiota in health and disease. Semin Immunol. 2013;25:370–7. https://doi.org/10.1016/j.smim.2013.09.005.

Scharschmidt TC, Fischbach MA. What lives on our skin: ecology, genomics and therapeutic opportunities of the skin microbiome. Drug Discovery Today: Disease Mechanisms. 2013;10:e83–9. https://doi.org/10.1016/j.ddmec.2012.12.003.

Schmidt E, Zillikens D. Pemphigoid diseases. Lancet. 2013;381:320–32. https://doi.org/10.1016/S0140- 6736(12)61140-4.

Schmidt U, Adan R, Böhm I, Campbell IC, Dingemans A, Ehrlich S, et al. Eating disorders: the big issue. Lancet Psychiatry. 2016;3:313–5. https://doi.org/10.1016/S2215-0366(16)00081-X.

Shen H, Rogelj S, Kieft TL. Sensitive, real-time PCR detects low-levels of contamination by Legionella pneumophila in commercial reagents. Mol Cell Probes. 2006;20:147–53. https://doi.org/10.1016/j.mcp.2005.09.007.

Slavich GM. Understanding inflammation, its regulation, and relevance for health: A top scientific and public priority. Brain Behav Immun. 2015;45:13–4. https://doi.org/10.1016/j.bbi.2014.10.012.

Słotwiński SM, Słotwiński R. Immune disorders in anorexia. Cent Eur J Immunol .2017;42:294–300. https://doi.org/10.5114/ceji.2017.70973.

Smith GD, Ebrahim S. Mendelian Randomization: Genetic Variants as Instruments for Strengthening Causal Inference in Observational Studies. In: National Research Council (US) Committee on Advances in Collecting and Utilizing Biological Indicators and Genetic Information in Social Science Surveys; Weinstein M, Vaupel JW, Wachter KW, editors. Biosocial Surveys. Washington (DC): National Academies Press (US); 2008.
16. Available from: https://www.ncbi.nlm.nih.gov/books/NBK62433/.

Song SJ, Lauber C, Costello EK, Lozupone CA, Humphrey G, Berg-Lyons D, et al. Cohabiting family members share microbiota with one another and with their dogs. ELife. 2013;2:e00458. https://doi.org/10.7554/eLife.00458.

Spencer CCA, Su Z, Donnelly P, Marchini J. Designing Genome-Wide Association Studies: Sample Size, Power, Imputation, and the Choice of Genotyping Chip. PLoS Genet. 2009;5:e1000477. https://doi.org/10.1371/journal.pgen.1000477.

Spor A, Koren O, Ley R. Unravelling the effects of the environment and host genotype on the gut microbiome. Nature Reviews Microbiology. 2011;9:279–90. https://doi.org/10.1038/nrmicro2540.

Srinivas G, Möller S, Wang J, Künzel S, Zillikens D, Baines JF, et al. Genome-wide mapping of gene– microbiota interactions in susceptibility to autoimmune skin blistering. Nat Commun. 2013;4:1–7. https://doi.org/10.1038/ncomms3462.

Stackebrandt E, Goebel. Taxonomic Note: A Place for DNA-DNA Reassociation and 16S rRNA Sequence Analysis in the Present Species Definition in Bacteriology. Int J Syst Evol Microbiol. 1994;44:846–9. https://doi.org/10.1099/00207713-44-4-846.

Starke R, Pylro VS, Morais DK. 16S rRNA Gene Copy Number Normalization Does Not Provide More Reliable Conclusions in Metataxonomic Surveys. Microb Ecol. 2021;81:535–9. https://doi.org/10.1007/s00248- 020-01586-7.

Steinhoff M, Neisius U, Ikoma A, Fartasch M, Heyer G, Skov PS, et al. Proteinase-Activated Receptor-2 Mediates Itch: A Novel Pathway for Pruritus in Human Skin. J Neurosci. 2003;23:6176–80. https://doi.org/10.1523/JNEUROSCI.23-15-06176.2003.

Stevens NE, Cowin AJ, Kopecki Z. Skin Barrier and Autoimmunity—Mechanisms and Novel Therapeutic Approaches for Autoimmune Blistering Diseases of the Skin. Front Immunol. 2019;10. https://doi.org/10.3389/fimmu.2019.01089.

Stewart JA, Chadwick VS, Murray A. Investigations into the influence of host genetics on the predominant eubacteria in the faecal microflora of children. J Med Microbiol. 2005;54:1239–42. https://doi.org/10.1099/jmm.0.46189-0.

Strumìa R, Varotti E, Manzato E, Gualandi M. Skin signs in anorexia nervosa. Dermatology. 2001;203:314–7. https://doi.org/10.1159/000051779.

Sung Kim H, Yosipovitch G. The Skin Microbiota and Itch: Is There a Link? J Clin Aesthet Dermatol. 2020;13:S39–46. https://doi.org/10.3390/jcm9041190.

Takeuchi O, Akira S. Pattern recognition receptors and inflammation. Cell. 2010;140:805–20. https://doi.org/10.1016/j.cell.2010.01.022.

Tremblay J, Singh K, Fern A, Kirton ES, He S, Woyke T, et al. Primer and platform effects on 16S rRNA tag sequencing. Frontiers in Microbiology. 2015;6. https://doi.org/10.3389/fmicb.2015.00771.

Turnbaugh PJ, Gordon JI. The core gut microbiome, energy balance and obesity. J Physiol. 2009;587:4153–8. https://doi.org/10.1113/jphysiol.2009.174136.

Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett C, Knight R, Gordon JI. The human microbiome project: exploring the microbial part of ourselves in a changing world. Nature. 2007;449:804–10. https://doi.org/10.1038/nature06244.

Uçkay I, Pittet D, Vaudaux P, Sax H, Lew D, Waldvogel F. Foreign body infections due to *Staphylococcus epidermidis*. Ann Med. 2009;41:109–19. https://doi.org/10.1080/07853890802337045.

Urashima R, Mihara M. Cutaneous nerves in atopic dermatitis. Virchows Archiv. 1998;432:363–70. https://doi.org/10.1007/s004280050179.

Vandegrift R, Bateman AC, Siemens KN, Nguyen M, Wilson HE, Green JL, et al. Cleanliness in context: reconciling hygiene with a modern microbial perspective. Microbiome. 2017;5:76. https://doi.org/10.1186/s40168-017-0294-2.

Walsh AM. Species classifier choice is a key consideration when analysing low-complexity food microbiome data Microbiome. 2018;6:50. https://doi.org/10.1186/s40168-018-0437-0

Wang J, Thingholm LB, Skiecevičienė J, Rausch P, Kummen M, Hov JR, et al. Genome-wide association analysis identifies variation in vitamin D receptor and other host factors influencing the gut microbiota. Nat Genet. 2016;48:1396–406. https://doi.org/10.1038/ng.3695.

Watanabe TK, Katagiri T, Suzuki M, Shimizu F, Fujiwara T, Kanemoto N, et al. Cloning and characterization of two novel human cDNAs (NELL1 and NELL2) encoding proteins with six EGF-like repeats. Genomics.
1996;38:273–6.

Wegiel B, Gallo D, Csizmadia E, Roger T, Kaczmarek E, Harris C, et al. Biliverdin inhibits Toll-like receptor-4 (TLR4) expression through nitric oxide-dependent nuclear translocation of biliverdin reductase. Proc Natl Acad Sci USA. 2011;108:18849. https://doi.org/10.1073/pnas.1108571108.

Westcott SL, Schloss PD. De novo clustering methods outperform reference-based methods for assigning 16S rRNA gene sequences to operational taxonomic units. PeerJ. 2015;3:e1487. https://doi.org/10.7717/peerj.1487.

Westmoreland P, Krantz MJ, Mehler PS. Medical Complications of Anorexia Nervosa and Bulimia. Am J Med. 2016;129:30–7. https://doi.org/10.1016/j.amjmed.2015.06.031.

Weyrich LS, Farrer AG, Eisenhofer R, Arriola LA, Young J, Selway CA, et al. Laboratory contamination over time during low-biomass sample analysis. Mol Ecol Resour. 2019;19:982–96. https://doi.org/10.1111/1755- 0998.13011.

Woese CR, Fox GE. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc Natl Acad Sci USA. 1977;74:5088–90. Woolston C. 'Does anyone have any of these?': Lab-supply shortages strike amid global pandemic. Nature. 2021. https://doi.org/10.1038/d41586-021-00613-y.

Wu GD, Chen J, Hoffmann C, Bittinger K, Chen Y-Y, Keilbaugh SA, et al. Linking Long-Term Dietary Patterns with Gut Microbial Enterotypes. Science. 2011;334:105–8. https://doi.org/10.1126/science.1208344.

Xu Z, Knight R. Dietary effects on human gut microbiome diversity. Br J Nutr. 2015;113:S1–5. https://doi.org/10.1017/S0007114514004127.

Yang B, Wang Y, Qian P-Y. Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. BMC Bioinformatics. 2016;17. https://doi.org/10.1186/s12859-016-0992-y.

Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, et al. Human gut microbiome viewed across age and geography. Nature. 2012;486:222–7. https://doi.org/10.1038/nature11053.

Zheng D, Liwinski T, Elinav E. Interaction between microbiota and immunity in health and disease. Cell Res. 2020;30:492–506. https://doi.org/10.1038/s41422-020-0332-7.

Zhou H, Shi L, Ren Y, Tan X, Liu W, Liu Z. Applications of Human Skin Microbiota in the Cutaneous Disorders for Ecology-Based Therapy. Front Cell Infect Microbiol. 2020;10. https://doi.org/10.3389/fcimb.2020.570261

Zhou Y, Zhi F. Lower Level of Bacteroides in the Gut Microbiota Is Associated with Inflammatory Bowel Disease: A Meta-Analysis. Biomed Res Int. 2016;2016:5828959. https://doi.org/10.1155/2016/5828959.

Zoetendal EG, Rajilic-Stojanovic M, de Vos WM. High-throughput diversity and functionality analysis of the gastrointestinal tract microbiota. Gut. 2008;57:1605–15. https://doi.org/10.1136/gut.2007.133603.

Zuo T, Ng SC. The Gut Microbiota in the Pathogenesis and Therapeutics of Inflammatory Bowel Disease. Front Microbiol. 2018;9:2247. https://doi.org/10.3389/fmicb.2018.02247.

Image references

1. Image by Mohammad2018.
   https://commons.wikimedia.org/wiki/File:Bullous_pemphigoid_new_image.jpg.
   Permitted for use under a CC BY-SA 4.0 license. https://creativecommons.org/licenses/by-sa/4.0/legalcode.

2. Image by Nouvelle iconographie de la Salpêtrière. Creator: Wellcome Library, London.
   Issued under the auspices of the Société de neurologie de Paris, 1906-18.
   https://commons.wikimedia.org/wiki/File:Louis_Bataille,_'Deux_cas_d'anorexie_hysterique'_Wellcom e_L0020548_(backcropped).jpg).
   Permitted for use under a CC BY-SA 4.0 license.
   https://creativecommons.org/licenses/by/4.0/legalcode.

3. Image by James Heilman, MD. "Atopy2010",
   https://commons.wikimedia.org/wiki/File:Atopy2010.JPG. Permitted for use under a CC BY-SA 4.0 license. https://creativecommons.org/licenses/by-sa/3.0/legalcode

Chapter 1

Characterization of the skin microbiota in bullous pemphigoid patients and controls reveals novel microbial indicators of disease

Belheouane, M.*, **Hermes, B. M.*,** Van Beek, N., Benoit, S., Bernard, P., Drenovska, K., Gerdes, S., Gläser, R., Goebeler, M., Günther, C., von Georg, A., Hammers, C. M., Holtsche, M. M., Homey, B., Horváth, O. N., Hübner, F., Linnemann, B., Joly, P., Márton, D., Patsatsi A., Pföhler C., Sárdy M., Huilaja L., Vassileva S., Zillikens D., Ibrahim S., Sadik C. D., Schmidt E., Baines, J. F. (2022). Characterization of the skin microbiota in bullous pemphigoid patients and controls reveals novel microbial indicators of disease. ***Journal of Advanced Research***. https://doi.org/10.1016/j.jare.2022.03.019

*Authors contributed equally to this work.

Journal of Advanced Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Advanced Research

journal homepage: www.elsevier.com/locate/jare

Original Article

Characterization of the skin microbiota in bullous pemphigoid patients and controls reveals novel microbial indicators of disease

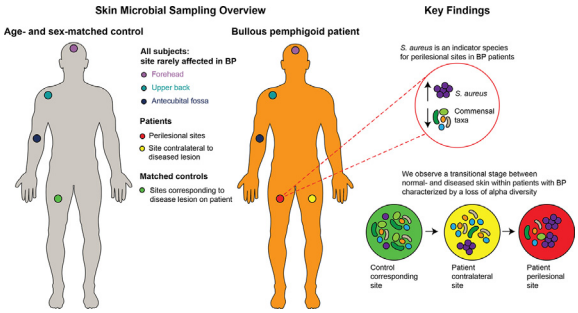
Meriem Belheouane [a,b,c,1], Britt M. Hermes [a,b,1], Nina Van Beek [d], Sandrine Benoit [e], Philippe Bernard [f], Kossara Drenovska [g], Sascha Gerdes [h], Regine Gläser [h], Matthias Goebeler [e], Claudia Günther [i], Anabelle von Georg [d], Christoph M. Hammers [d,r], Maike M. Holtsche [d], Bernhard Homey [j], Orsolya N. Horváth [k], Franziska Hübner [d], Beke Linnemann [d], Pascal Joly [l], Dalma Márton [m], Aikaterini Patsatsi [n], Claudia Pföhler [o], Miklós Sárdy [k,m], Laura Huilaja [p], Snejina Vassileva [g], Detlef Zillikens [d,q], Saleh Ibrahim [q,r,s], Christian D. Sadik [d,q], Enno Schmidt [d,q,r,1], John F. Baines [a,b,1,*]

[a] *Max Planck Institute for Evolutionary Biology, August-Thienemann-Str. 2, 24306 Plön, Germany*
[b] *Section of Evolutionary Medicine, Institute for Experimental Medicine, Kiel University, Arnold-Heller-Str. 3, 24105 Kiel, Germany*
[c] *Research Center Borstel, Evolution of the Resistome, Leibniz Lung Center, Parkallee 1-40, 23845 Borstel, Germany*
[d] *Department of Dermatology, Allergy, and Venereology, University of Lübeck, 23538 Lübeck, Germany*
[e] *Department of Dermatology, Venereology and Allergology, University Hospital Würzburg, Würzburg, Germany*
[f] *Department of Dermatology, Reims University Hospital, University of Reims Champagne-Ardenne, Reims, France*
[g] *Department of Dermatology and Venereology, Medical University - Sofia, 1, St. G. Sofiiski str, 1431 Sofia, Bulgaria*
[h] *Department of Dermatology, Venereology and Allergology, University of Kiel, 24105 Kiel, Germany*
[i] *Department of Dermatology, University Hospital, TU Dresden, 01307 Dresden, Germany*
[j] *Department of Dermatology, University Hospital Düsseldorf, Medical Faculty, Heinrich-Heine-University Düsseldorf, Moorenstr. 5, 40225 Düsseldorf, Germany*
[k] *Department of Dermatology and Allergy, University Hospital, LMU Munich, Munich, Germany*
[l] *Department of Dermatology, Rouen University Hospital, INSERM U1234, Normandie University, Rouen, France*
[m] *Department of Dermatology, Venereology and Dermatooncology, Semmelweis University, Budapest, Hungary*
[n] *Autoimmune Bullous Diseases Unit, 2nd Dermatology Department, Aristotle University School of Medicine, Papageorgiou General Hospital, 56403 Thessaloniki, Greece*
[o] *Department of Dermatology, University of Saarland, Homburg/ Saar, Germany*
[p] *PEDEGO Research Unit, University of Oulu; Department of Dermatology and Medical Research Center Oulu, Oulu University Hospital, Oulu, Finland*
[q] *Center for Research on Inflammation of the Skin (CRIS), University of Lübeck, 23538 Lübeck, Germany*
[r] *Lübeck Institute of Experimental Dermatology (LIED), University of Lübeck, Lübeck, Germany*
[s] *College of Medicine and Sharjah Institute for Medical Research, University of Sharjah, 27272 Sharjah, UAE*

HIGHLIGHTS

- We conducted a large-scale investigation of skin microbiota composition and diversity in BP.
- We reveal substantial differences in skin microbiota in patients with BP compared to that of control subjects.
- We observe a transitional stage between normal- and diseased skin within patients with BP.
- BP is characterized by a loss of protective microbiota and an increase in *S. aureus*, an inflammation-promoting species.

GRAPHICAL ABSTRACT



Abbreviations: AI, autoimmune; BP, bullous pemphigoid; AIBD, Autoimmune blistering disease; BPDAI, Bullous Pemphigoid Disease Area Index; DF, degrees of freedom; IS, indicator species; ASV, amplicon sequence variant.

Peer review under responsibility of Cairo University.

* Corresponding author at: Max Planck Institute for Evolutionary Biology, August-Thienemann-Str. 2, 24306 Plön, Germany.
E-mail addresses: baines@evolbio.mpg.de, baines@mpg.evolbio.de (J.F. Baines).
[1] Authors contributed equally to this work.

https://doi.org/10.1016/j.jare.2022.03.019
2090-1232/

- *S. aureus* ubiquitously associates with BP, suggesting a role in pathogenesis.
- Our results may help inform clinical markers for assessing BP disease risk and prognosis.

ARTICLE INFO

Article history:
Received 13 December 2021
Revised 27 March 2022
Accepted 31 March 2022
Available online xxxx

Keywords:
Autoimmune blistering disease
Bullous pemphigoid
Skin microbiota
16s rRNA gene sequencing
Risk factor

ABSTRACT

Introduction: Bullous pemphigoid (BP) is the most common autoimmune blistering disease. It predominately afflicts the elderly and is significantly associated with increased mortality. The observation of age-dependent changes in the skin microbiota as well as its involvement in other inflammatory skin disorders suggests that skin microbiota may play a role in the emergence of BP blistering. We hypothesize that changes in microbial diversity associated with BP might occur before the emergence of disease lesions, and thus could represent an early indicator of blistering risk.
Objectives: The present study aims to investigate potential relationships between skin microbiota and BP and elaborate on important changes in microbial diversity associated with blistering in BP.
Methods: The study consisted of an extensive sampling effort of the skin microbiota in patients with BP and age- and sex-matched controls to analyze whether intra-individual, body site, and/or geographical variation correlate with changes in skin microbial composition in BP and/or blistering status.
Results: We find significant differences in the skin microbiota of patients with BP compared to that of controls, and moreover that disease status rather than skin biogeography (body site) governs skin microbiota composition in patients with BP. Our data reveal a discernible transition between normal skin and the skin surrounding BP lesions, which is characterized by a loss of protective microbiota and an increase in sequences matching Staphylococcus aureus, a known inflammation-promoting species. Notably, Staphylococcus aureus is ubiquitously associated with BP disease status, regardless of the presence of blisters.
Conclusion: The present study suggests Staphylococcus aureus may be a key taxon associated with BP disease status. Importantly, we however find contrasting patterns in the relative abundances of Staphylococcus hominis and Staphylococcus aureus reliably discriminate between patients with BP and matched controls. This may serve as valuable information for assessing blistering risk and treatment outcomes in a clinical setting.

Introduction

Bullous pemphigoid (BP) is the most common autoimmune skin blistering disease (AIBD), with an annual incidence of about 20 new cases per million in Europe [1–3]. It occurs when autoantibodies attack two structural hemidesmosomal proteins of the epidermal basement membrane, i.e., BP180 (type XVII collagen) and BP230, resulting in subepidermal blistering [2,4,5]. The severity of this highly pruritic AIBD considerably affects quality of life and is associated with significantly increased mortality [5]. The incidence of BP is increasing with the aging European population [2,3]. It is thus intriguing that a recent multinational study of 9,000 participants showed that skin microbiota are a predictor of age, more so than oral or gut microbiota [6].

The observation of age-dependent changes in the skin microbiota as well as its involvement in other inflammatory skin disorders suggest that skin microbiota may play a role in the emergence of AIBD [7–13]. While certain HLA haplotypes are associated with BP, such as HLA-DQA1*05:05 and HLA-DRB1*07:01 in Germans [14], few triggering factors apart from age, medication use, and neuro-psychiatric disease are described [2,15,16]. Previous efforts by Srinivas et al. [17] demonstrated that genotype-dependent microbiota affect disease susceptibility in a mouse model of epidermolysis bullosa acquisita, an AIBD with autoantibodies against type VII collagen. Similarly, Miodovnik et al. [18] presented pilot human data suggesting that skin microbiota contribute to the pathogenesis of BP. However, identification of candidate bacterial taxa or important changes in microbial diversity associated with BP remain uncharacterized.

Here, we conducted a large-scale investigation of patients with BP and age- and sex-matched controls within Europe to clarify relationships between microbiota and BP. By examining skin microbiota surrounding (i) BP lesions, (ii) unaffected skin areas in BP patients, and (iii) controls matched for sex, age, and body site, we reveal clear microbial indicators of both BP disease status and blistering status. The detection of microbial taxa associated with AIBD blistering could enable early intervention and thus, better clinical outcomes.

Materials and methods

Ethics statement

All experiments involving human subjects were conducted according to the ethical policies and procedures approved by the ethics committee of the University of Lübeck (Approval no. 15–051, 18–046), as well as the respective committees of the study centers, following the Declaration of Helsinki. Written, informed consent was obtained from each participant.

Study participants

Four-hundred eighteen volunteers were recruited from fourteen study centers across Europe (Germany: Dresden, Düsseldorf, Freiburg, Homburg, Kiel, Lübeck, Munich, Würzburg; France: Reims, Rouen; Sofia, Bulgaria; Thessaloniki, Greece; Oulu, Finland) between October 2015 and September 2019. All volunteers were of European descent. Patients with BP were diagnosed according to national and international guidelines and had (i) a compatible clinical picture, (ii) linear deposits of IgG and/or C3 along the dermal-epidermal junction by direct immunofluorescence of a perilesional

skin biopsy, and iii) serum IgG reactivity against the epidermal side of human salt-split skin or BP180 NC16A ELISA [19,20]. Patients with BP (n = 228) included 114 males, 113 females, and one sex "unspecified" participant, with an average age of 80 ± 8.95 (SD) years (range, 49 to 98 years), and with newly diagnosed or relapsed BP. No newly diagnosed patient had received systemic treatment using dapsone, doxycycline, or immunosuppressants (with the exception of corticosteroids described below) at the time of sampling. All patients had abstained from topical antiseptics two weeks prior to sampling. Systemic and topical corticosteroids had not been administered for BP for longer than 7 days before skin swaps were taken. None of the swabbed individuals received antibiotic therapy, including doxycycline, for at least four weeks.

Age- and sex-matched controls (n = 190) included 104 males and 86 females with non-inflammatory/non-infectious dermatoses with an average age of 80 years ± 8.51 (SD) years (range, 47 to 100 years). Controls did not receive systemic antibiotics for at least four weeks prior to sampling. Clinical metadata used for the analysis are provided in Supplementary Table S1; summarized demographic and clinical data are provided in Supplementary Table S2. Pictures showing normal skin and a typical BP lesion of a patient included in our study are provided in Supplementary Figure 1.

Sampling, DNA extraction, and 16S rRNA gene sequencing

Samples were collected using Epicentre Illumina collection swabs (Madison, WI, USA) immersed in 600 uL buffer (50 mM Tris, 1 mM EDTA, 0.5% Tween-20) (Teknova, United States). The swabs were rubbed across the selected body site for 30 s and then placed back into the buffer solution. Immediately after swabbing, swabs were stored at −80 °C until further processing.

Skin samples (n = 2,956) were obtained from patients with BP representing different cutaneous microenvironments, including "perilesional" skin (defined as being within 2 cm of a primary BP lesion, i.e. a fresh blister or erosion), unaffected skin at the same anatomical location on the contralateral side of the patient (referred to as "contralateral"), and unaffected skin in areas that do not typically manifest disease (we selected the forehead and upper back, as described by Schmidt and Groves [21]), in addition to the antecubital fossa, which was sampled in the human microbiome project [22], collectively referred to as "sites rarely affected by BP"). Two separate perilesional sites from anatomically different BP lesions were sampled from each patient to account for differences in skin biogeography. The locations of lesioned skin, and therefore, perilesional sampling sites, varied from patient to patient. The most common sites included the thigh, arm, foot, knee, lower leg, and hand.

Control participants were swabbed at locations that approximated the sampled body sites in the patients with BP (referred to as "corresponding sites"), in addition to the three sites rarely affected by BP (Fig. 1a). Ambient air samples (n = 19), collected by holding a swab in the air for 30 s and then placing the collection swab directly into the buffer solution, served as negative sampling controls in addition to negative extraction controls (n = 43). Negative controls were processed alongside samples.

ZymoBIOMICS Microbial Community Standard cells (Zymo Research) were used as extraction and sequencing controls to assess contamination in downstream analyses, following the mock community dilution series protocol as described by Karstens et al. [23]. In brief, the strategy is based on the logic that with decreasing "true" microbial biomass (i.e., skin microbes or mock cells), potential signal from background/contamination introduced throughout the procedure will increase. All mock dilutions, as well as the undiluted mock community standards, were treated as samples throughout the extraction, PCR, sequencing, and data processing steps.

Swabs immersed in buffer were thawed overnight at 4℃, then vortexed at high speed for 1 min. After swab removal, tubes were centrifuged at 13,000 g for 15 min, and the pellets were resuspended in Power Bead solution. DNA was subsequently extracted using the Qiagen DNeasy UltraClean 96 Microbial Kit [96-well plate] (Germantown, MD, USA), according to the manufacturer's instructions, and eluted in 50 uL of the elution buffer. Negative extraction controls were included for each 96-well plate. Samples were stored at −20℃ until further processing.

PCR and sequencing were performed by implementing the dual-index sequencing strategy for amplicon sequencing on the MiSeq Illumina sequencing platform, as previously described [24]. Final sample sizes included 2,319 skin swabs comprising 1,451 patient and 868 matched control swabs. A detailed description of sampling methodology and sample processing is provided in Supplementary Methods.

Data processing and taxonomic classification

The challenges of studying low biomass communities such as skin microbiota are well-documented and include exogenous bacterial DNA contamination from sources such as laboratory reagents, air, and sample collection instruments [23,25–28]. A detailed description of the steps implemented to account for potential contamination is provided in Supplementary Methods. Briefly, data processing and statistical analyses were performed using R (version 4.0.2). Sequences were processed using DADA2 (version 1.16.0), resulting in abundance tables of amplicon sequence variants (ASVs) [29]. To normalize sequencing coverage, random sub-sampling to 5,000 sequences per sample was performed [30]. Decontam (version 1.8.0; [25] was used within Phyloseq (version 1.32.0) [31] to identify potential contaminant ASVs, according to the prevalence method [23]. ASVs classified to families Halomonadaceae (n = 1,040) and Shewanellaceae (n = 211) were removed, following recommendations of Weyrich et al. [32]. Summary read data are provided in Supplementary Table S3. Taxonomic assignment of ASVs was completed in DADA2 with the Bayesian classifier using the NR Silva database training set, version 138 [33]. Representative 16S rRNA gene sequences were queried via the Ribosomal Database Project (RDP; release 11.6; [34] SeqMatch (version 3; [35]; Supplementary Table S4).

Ecological and statistical analyses

Analyses included several patient and disease categories. Disease status refers to patients with BP versus matched controls. Blistering status refers to patient perilesional sites versus unaffected, contralateral sites of the same patient. Disease activity was calculated using the Bullous Pemphigoid Disease Area Index (BPDAI) [36]. The activity score of both skin and mucosa were combined to account for disease activity, while damage and pruritus points were not considered for calculations. Supplementary Table S5 provides the BPDAI scores for study participants.

Statistical analyses were performed in R (version 4.0.2). Alpha diversity was measured using Shannon and Chao1 indices with vegan (version 2.5–6) on absolute abundance data. Beta diversity was calculated using the Bray-Curtis dissimilarity index. We performed a non-parametric multivariable analysis of variance using distance matrices (PERMANOVA) using the "adonis" function with 1,000 permutations and a partial constrained principal coordinate analysis of beta diversity measures using the "capscale" function in vegan [37]. The significance of models, axes, and terms were assessed using the "anova.cca" function with 1,000 permutations.

Indicator species analysis was applied using indicspecies (version 1.7.9) with the "r.g." function [38] and 100,000 permutations. Random Forest classification and regression analyses were performed using randomForest (version 4–6-14) [39]. Models were constructed with 100,000 trees, with the "mtry" parameter set for

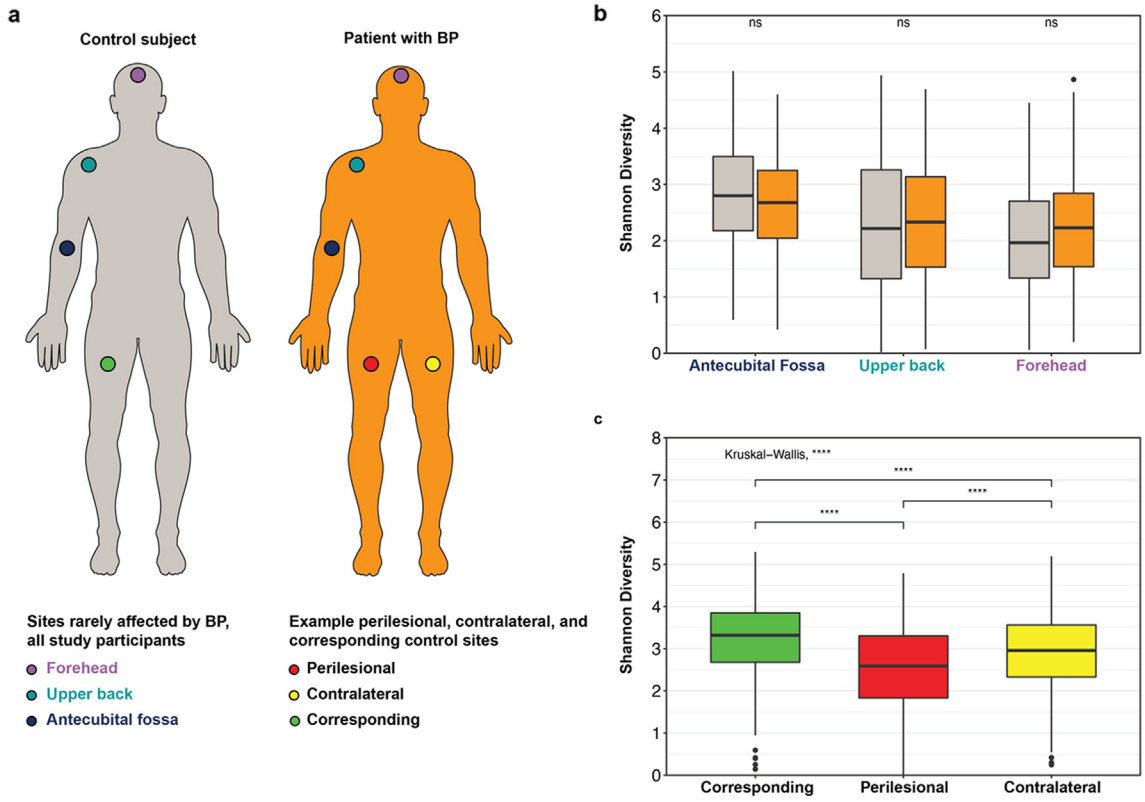


Fig. 1. Sampling sites for patients with BP and matched controls; Box plots of Shannon (alpha) diversity. 1a. Grey figure represents age-and sex-matched control; orange figure represents a patient with BP. Sites rarely affected by BP [2] include the forehead (purple), upper back (turquoise), and antecubital fossa (dark blue) are represented on both figures. An example perilesional sampling site (red), unaffected contralateral site (yellow) on the patient, and control-matched corresponding site (green) are shown. 1b. Shannon diversity at the ASV-level for sites rarely affected by BP for controls and patients. 1c. Shannon diversity at the ASV-level for patient perilesional, patient contralateral, and control corresponding sites. For box plots: Boxes represent interquartile range between first and third quartiles; horizontal line defines the median. Whiskers represent the 5th and 95th percentiles and values beyond these bounds are considered outliers, marked with black dots. Kruskal-Wallis test applied to analyze site variation. If an overall significant difference was observed, a pairwise Wilcox test was performed; p-values adjusted using the Benjamini-Hochberg method. Significance represented by: * ≤ 0.05; ** ≤ 0.01; *** ≤ 0.001; **** ≤ 0.0001; ns = not significant. Supplementary Table S6 reports summary statistics.

each model and linear models constructed to evaluate potential disease effects. Adjusted R^2 values reported, beta coefficient values express directionality. Further details are provided in Supplementary Methods.

Results

Sampling

Two hundred twenty-eight patients with BP and 190 age- and sex-matched controls from fourteen study sites across Europe were included (see Methods). We performed 16S rRNA gene sequencing on bacterial genomic DNA derived from swabbing four categories of body sites (Fig. 1a). These include areas adjacent to a fresh blister or erosion (“perilesional”), non-lesional skin contralateral to the perilesional sample on the same patient (“contralateral”), and the same body site on an age- and sex-matched control (“corresponding”). The locations of the perilesional sites varied from patient to patient. Sites considered to be rarely affected in BP, i.e., forehead and upper back [21], as well as the antecubital fossa, which was sampled in the human microbiome project [22], were sampled in both patients with BP and controls to obtain a more complete picture of the skin microbiota in BP across skin biogeography (Fig. 1a).

Reduced alpha diversity within lesional and BP-susceptible sites

To assess alpha diversity, we included both the Shannon and Chao1 indices, which reflect taxon evenness and richness, respectively. At the ASV-level, we found that the Shannon (Fig. 1b) and Chao1 (Supplementary Figure S2a) indices are similar in patients and controls at sites rarely affected by BP. In contrast, control corresponding sites display higher bacterial diversity than patient contralateral sites, which in turn are more diverse than perilesional sites for Shannon (Fig. 1c) and Chao1 indices (Supplementary Figure S2b). Supplementary Table S6 provides the summary statistics for group comparisons.

Critically, study center, disease status (i.e., patient with BP versus matched control), and sex significantly correlate with Shannon diversity for patient perilesional and contralateral sites as well as for control corresponding sites ($F_{37,1118} = 7.24$; R^2_{adj}: 0.17; $p < 0.001$), with disease status explaining 8.28% of the variance and study center and sex explaining 5.3% and 1.3% of the variance, respectively. Likewise, disease status and study center significantly correlate with Chao1 richness ($F_{37,1118} = 6.03$; $R^2_{adj} = 0.14$; $p < 0.001$), with study center explaining 7.86% and disease status explaining 1.41% of the Chao1 variance. Disease status associates with a decrease in Shannon diversity in patient perilesional and contralateral sites ($\beta = -0.72, -0.38$, respectively), and a decrease in Chao1 richness ($\beta = -39.38, -30.90$, respectively). Thus, disease status associates with a substantial decrease in both Shannon diversity and Chao1 richness, which is still present after accounting for potential confounding variables. To determine whether these findings were affected by spatial correlation across body sites, we calculated a linear mixed model using “individual” as a random term to estimate variability in alpha diversity measures and to control for non-disease variables, including sex and study center. The model reveals statistically significant variance similar to that estimated

by the above linear models, suggesting that the results reported here are unlikely to be conflated by cases of multiple measures of diversity (see Supplementary Results).

Analysis of sum of squares shows that the effect of disease status on alpha diversity does not extend to sites rarely affected by BP. Rather, skin biogeography likely characterizes microbial diversity at these sites (see Supplementary Results). Of note, the severity of disease as determined by the Bullous Pemphigoid Disease Area Index (BPDAI) [36], does not significantly associate with mean alpha diversity measures at perilesional and contralateral skin sites.

Beta diversity in relation to disease, individual, and sampling features

We first analyzed beta (between-sample) diversity at sites rarely affected by BP to evaluate the effects of potential confounding variables (see Supplementary Methods). Analysis of disease status per body site reveals a significant association with disease status (adonis: Bray-Curtis ~ disease status: body site; $R^2 = 0.003$; $p < 0.001$). Partial constrained principal coordinate analysis reveals that patients and controls cluster according to body site along the first and second axes and that forehead and upper back (typical sebaceous zones) are more similar to each other compared to the antecubital fossa (Fig. 2b). These findings suggest that the microbial variation among sites rarely affected by BP is likely linked to skin biogeography rather than disease or study center (Fig. 2a, 2c). Additionally, partial constrained principal coordinate analysis reveals that on the first and second axes, control corresponding sites are distinguishable from patient contralateral and perilesional sites, which largely cluster together (Fig. 2d, e). Fig. 2f, on the other hand, shows comparatively little clustering according to study center.

We additionally analyzed beta diversity between patient perilesional, patient contralateral, and control corresponding sites as described above. We find that disease status, blistering status (patient perilesional sites versus unaffected, contralateral sites of the same patient), and study center all explain a portion of the variance in beta diversity (see Supplementary Results). However, an analysis of interaction between variables reveals that disease status accounts for significant differences between study centers (adonis: disease status: study center, $R^2 = 0.03$; $p < 0.001$). Furthermore, linear modeling shows that BPDAI (i.e., disease severity) significantly correlates with study center ($F_{13,196} = 3.31$, $R^2_{adj} = 0.117$, $p < 0.001$). These results suggest that variation between study centers could be explained by differences in patient populations between study centers, e.g., perhaps only the most severely affected patients are seen at university study centers in some regions.

Indicator species of BP patients and controls

We conducted four indicator species analyses at the ASV-level. To refine the taxonomic classification of indicator ASVs, we queried representative sequences using RDP SeqMatch (see Supplementary Table S4). ASVs strongly associated with BP patients or controls are shown in Fig. 3 and Supplementary Table S7.

Several indicator ASVs known to be human commensals associate with sites rarely affected by BP (i.e., forehead, upper back, and antecubital fossa in our study). Importantly, we identify a greater number of ASVs associating with these standardized control sites, which is coherent with the observed loss of diversity in patients with BP. ASV_1, which closely matches *Cutibacterium acnes* (*C. acnes*) [previously known as *Propionibacterium acnes;* [40], is an indicator at the forehead, upper back, and antecubital fossa. Specifically, among these sites rarely affected by BP, body site is associated with 13% of the variance in *C. acnes* abundance (abundance is accordingly higher at the sebaceous forehead and upper back sites), whereas disease status accounts for 0.79% variance of *C. acnes* abundance ($F_{39,1123} = 11.58$; $R^2_{adj} = 0.26$; $p < 0.001$). Interestingly, disease status is associated with a decrease in *C. acnes*

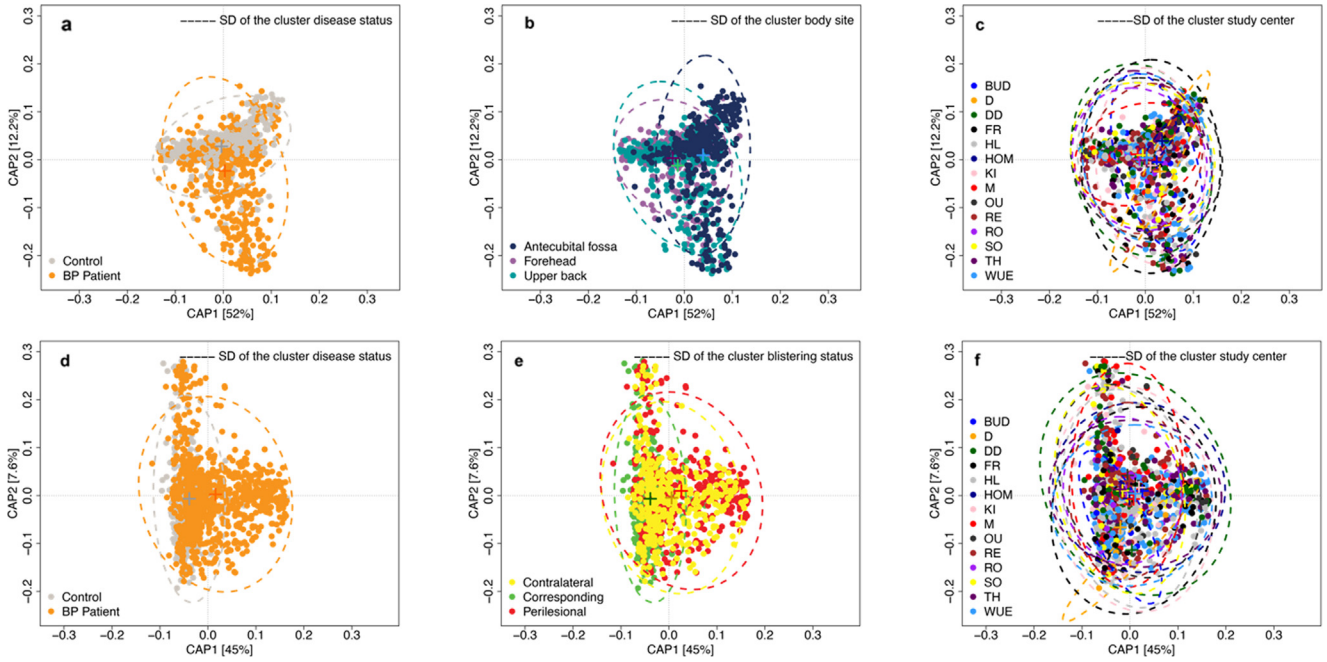


Fig. 2. Partial constrained principal coordinate analyses of Bray-Curtis 2a to 2c. Body sites rarely affected by BP. (anova.cca, Full model: $p = 0.0009$; terms: disease status, body site (constrained inertia = 5.04%, conditioned inertia = 4.5%), study center: $p < 0.001$; axes: CAP1, CAP2: $p = 0.09$; 1,000 permutations). 2d to 2f. Patient perilesional, patient contralateral sites, and control corresponding sites. (anova.cca, Full model: $p < 0.001$; terms: disease status, blistering status, study center: $p < 0.001$; axes: CAP1, CAP2: $p = 0.009$; 1,000 permutations; see "ecological and statistical analysis" in Methods). "+" represents centroid. SD: standard deviation. Site abbreviations: Budapest, Hungary (BUD); Düsseldorf, (D), Dresden, (DD), Freiburg, (FR), Lübeck, (HL), Homburg, (HOM), Kiel, (KI), München, (M), Würzburg (WUE), all Germany; Oulu, Finland (OU); Reims, (RE), Rouen, (RO), both France, Sofia, Bulgaria (SO), Thessaloniki, Greece, (TH).

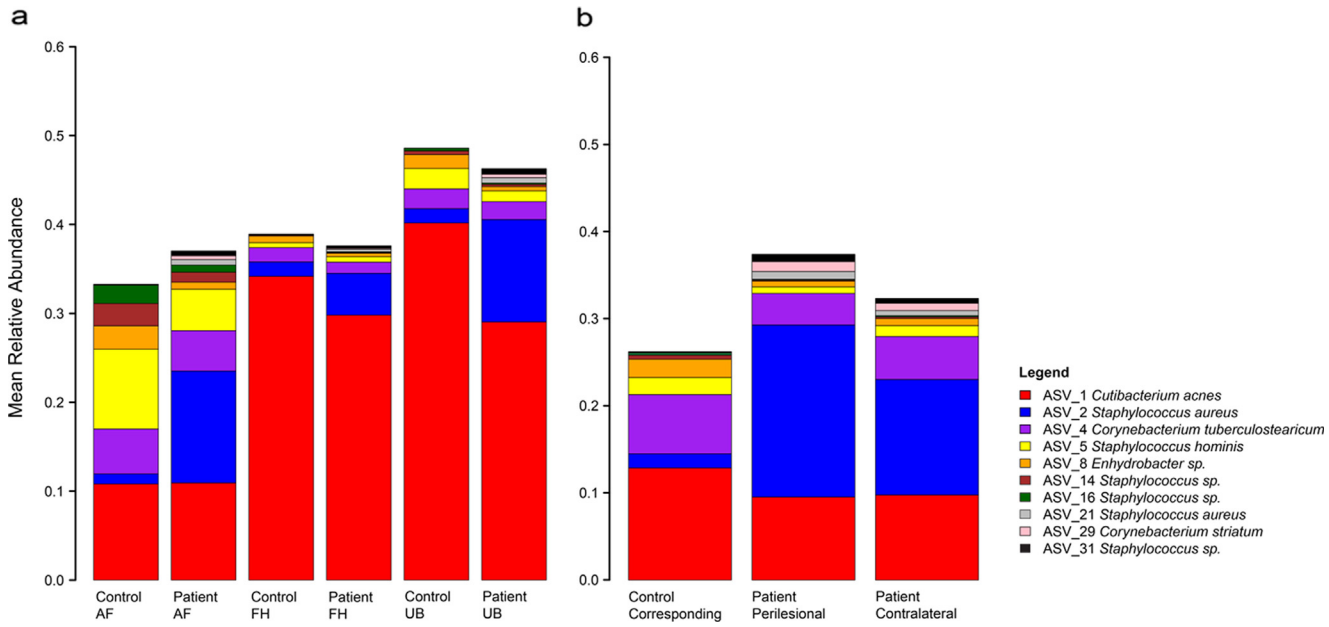


Fig. 3. Bar plots of mean relative abundance for the ten most important indicator species. 3a. Bar plot showing relative abundance of important indicator species, at the ASV-level, for controls and patients with BP at sites rarely affected by BP [antecubital fossa (AF), forehead (FH), and upper back (UB)]. 3b. Bar plot showing the relative abundances of important indicator species at the ASV-level for patient perilesional, patient contralateral sites, and control-matched corresponding sites. RDP SeqMatch results for the representative ASV sequences are shown in the legend and provided in full in Supplementary Table S4. Supplementary Tables S7, S8 provide statistical parameters for indicator species analyses and summary statistics of all indicator ASVs, respectively.

abundance at the upper back and forehead (β = –0.11 and β = –0.045, respectively).

Important patterns are also apparent among control corresponding, patient perilesional, and patient contralateral sites. Within these sites, *C. acnes* abundance associates with study center, blistering status, and sex ($F_{37,1118}$ = 5.40; R^2_{adj} = 0.12, $p < 0.001$), with sex explaining 7.17%, study center explaining 4.66%, and blistering status (i.e., patient perilesional versus patient contralateral) explaining 0.63% of the variance. Furthermore, *C. acnes* relative abundance is greater at control corresponding sites and relatively lower at patient perilesional sites (β = +0.03, –0.003, respectively). The higher relative abundance of *C. acnes* at these control corresponding sites is consistent with the increased abundance of *C. acnes* observed in rarely affected sites. Additionally, ASV_4 [which closely matches *Corynebacterium tuberculostearicum* (*C. tuberculostearicum*)] is an indicator for control corresponding sites and patient contralateral sites. Study center, blistering status, and sex are significantly associated with *C. tuberculostearicum* abundance ($F_{37,1188}$ = 3.92; R^2_{adj} = 0.09; $p < 0.001$), with blistering status explaining 4.06% of the variance and correlating with an increase in abundance in control corresponding sites, but a decrease in patient perilesional sites (β = +0.05, –0.04, respectively). Summary statistics for indicator ASVs are provided in Supplementary Table S8.

Contrasting patterns of Staphylococcus ASVs in patients with BP and controls

Six indicator ASVs belong to *Staphylococcus* and display contrasting patterns associated with disease status (i.e., patient with BP versus matched control). *Staphylococcus* ASV_5 (which closely matches *Staphylococcus hominis* [*S. hominis*]) abundance significantly correlates with both disease status and body site ($F_{39,1123}$ = 6.45; R^2_{adj} = 0.16; $p < 0.001$). However, as with other indicator ASVs known to be human commensals, body site explains a greater proportion of variance (11.17%) compared to disease status (1.24%), whereby the latter is associated with a decrease in abundance (β = –0.04). Accordingly, *S. hominis* is significantly negatively correlated with BPDAI (i.e., disease severity) at patient contralateral sites (Spearman's rho = -0.17; $p < 0.05$), but there is no relationship between BPDAI and *S. hominis* at patient perilesional sites. In contrast, *Staphylococcus* ASV_2 (which closely matches *Staphylococcus aureus* [*S. aureus*]) is a strong indicator for BP, including at patient body sites rarely affected by BP. Here, disease status explains 7.35% of the variance in *S. aureus* abundance, whereas body site explains 1.52% ($F_{39,1123}$ = 5.12; R^2_{adj} = 0.12; $p < 0.001$). Notably, disease status associates with an increase in *S. aureus* abundance at these rarely affected sites (β = 0.08). However, among perilesional, contralateral, and corresponding sites, blistering status (patient perilesional versus unaffected, contralateral sites of the same patient), accounts for the greatest amount of variance for *S. aureus* abundance (11.67%), followed by study center (6.45%), and sex (0.77%; $F_{37,1118}$ = 9.09; R^2_{adj} = 0.21, $p < 0.001$). This is characterized by a decrease of *S. aureus* abundance in control corresponding sites (β = –0.12) compared to an increase at patient perilesional sites (β = +0.07). Additionally, *S. aureus* positively correlates with BPDAI at perilesional and contralateral sites (Spearman's rho = 0.2; $p < 0.01$; Spearman's rho = 0.28; $p < 0.001$, respectively; Supplementary Figure S3). To address concerns of spatial correlation across body sites as a potential confounding factor, we constructed a linear mixed model using "individual" as a random term to control for non-disease variables and to estimate variability in mean ASV indicator abundances explained by BPDAI for patients with BP. Estimates reveal similar findings in terms of significance and proportions of variance explained by disease status, except for ASV_1, *C. acnes*, which is not affected by blistering status using this model (Supplementary Results).

Because individual members of *Staphylococcus* can display antagonistic interactions in the context of inflammatory skin disorders [41], we examined pairwise correlations among the top ten indicator ASVs (Supplementary Figures S4, S5; Supplementary Table S9). Importantly, *Staphylococcus* ASV_2 (*S. aureus*) and ASV_5 (*S. hominis*) display significant negative correlations within patient perilesional sites, patient contralateral sites, and at the antecubital fossa site in patients with BP. However, there is no

significant correlation between these two *Staphylococcus* indicators at any matched control sites. Furthermore, *Staphylococcus* ASV_2 is significantly negatively correlated with sequences matching *C. acnes* (ASV_1) at all sampling category sites in patients with BP. This association is absent at all sampling category sites in matched controls. This finding suggests a fundamental alteration in community interactions among members of *Staphylococcus* in the context of BP.

Finally, because BP is an age-related disease and the skin microbiota is also generally known to display age-dependent changes, we evaluated whether *S. aureus* itself may display age-dependent change, e.g., an increase in abundance with age. However, we do not find ASV_2 abundance to correlate with age at either perilesional (Spearman's rho = 0.687; $p = 0.32$) or at contralateral sites (Spearman's rho = 0.05; $p = 0.43$).

Staphylococcus ASVs predict disease status in random forest classification

Random forest classification analyses reveal indicator ASVs to accurately classify samples when applied to all sampling sites (mtry = 15; 849/868 controls and 1,443/1,451 BP patients; mean classification accuracy 99.00%). Prediction accuracy approaches 100% when applied to only control corresponding, patient contralateral, and patient perilesional sites (mtry = 18; 324/334 controls and 822/822 BP patients; mean classification accuracy 99.15%; Supplementary Figures S6a, S6b). By inspecting the mean decrease accuracy components for ASVs, those belonging to the *Staphylococcus* genus are identified as being most important to both models (see Supplementary Table S10).

To estimate the discriminatory power of *Staphylococcus* ASVs alone, we limited the random forest classification analyses to *Staphylococcus* ASVs with an abundance greater than 2% within each sample. We found that *Staphylococcus* ASVs accurately distinguish between controls and patients with BP (mtry = 52; 790/868 controls and 1,446/1,451 patients; mean classification accuracy 96.40%) when applied to all sampling sites and are similarly accurate when applied using only control corresponding, patient perilesional, and patient contralateral sites (mtry = 62; controls 294/334 and 819/822 patients; 96.20%; Supplementary Figures S6c, S6d). Notably, inspection of mean decrease accuracy components indicates that *S. aureus* ASV_2 is the most important ASV for model accuracy.

Discussion

This study reveals marked differences in the skin microbiota of patients with BP compared to that of sex- and age-matched controls with non-inflammatory/ non-infectious dermatoses. This was accomplished by conducting large-scale sampling and bacterial 16S rRNA gene analysis, utilizing a sampling scheme that accounts for both skin biogeography and disease status (Fig. 1a). This study represents the most substantial sampling effort of skin microbiota in BP to date.

We observe a significant reduction in alpha diversity at both perilesional sites and contralateral sites in BP patients compared to site-matched areas from controls. Furthermore, blistering status (i.e., patient perilesional sites), as well as disease status (i.e., patients with BP disease versus controls), are associated with a fewer number of indicator ASVs when compared to matched corresponding sites from control subjects. This reduction in alpha diversity in patients with BP is consistent with findings from other studies of inflammatory skin diseases, including psoriasis [8,9], atopic dermatitis [42], as well as a mouse model of the BP-like variant epidermolysis bullosa acquisita [17].

The clear biogeography of human skin microbiota, whereby distinct assemblages colonize different body sites depending upon numerous factors, suggests that conditions like BP that affect the skin micro-environment, and thereby skin microbiota, may influence susceptibility to blistering [26,43]. Our data reveal that BP might contribute to a loss of protective microbiota in sites rarely affected by BP. At the upper back, an interaction model revealed that disease status significantly associates with a decrease in *C. acnes* relative abundance in patients with BP. This is notable given that the upper back represents a sebaceous skin zone where we would expect relatively high amounts of *C. acnes* [13,44]. Although *C. acnes* is commonly thought of as a potential pathogenic species responsible for acne, it also acts as an important commensal that aids in preventing the colonization and invasion of pathogens via the production of antimicrobials and hydrolysis of triglycerides [13,45], as well as the production of short-chain fatty acids [46,47]. For *S. hominis,* another human commensal, we also find that disease status associates with a decrease in abundance at rarely affected sites. Furthermore, we find a negative association between disease activity (measured by the validated disease score BPDAI [36]) and *S. hominis* in contralateral sites of patients with BP. Additionally, our data show that in the skin sites of matched controls that correspond to the perilesional sites in patients with BP, there is a relative increase in abundance of the commensals *C. acnes* and *C. tuberculostearicum*. Furthermore, *S. hominis* is relatively decreased in patient contralateral sites, suggesting that the effect of disease extends beyond perilesional sites in patients with BP. It is thus possible that protective effects provided by different commensal bacteria may be fundamentally altered in patients with BP, e.g., the production of antimicrobials by *Staphylococcus* strains, as in atopic dermatitis [41,48]. Decreased commensal microbiota perhaps translate to fewer protective immune functions in the skin, which in turn could allow for increased colonization of inflammation-promoting species like *S. aureus* [41]. Thus, our observations might be capturing a baseline state of disease at sites without blisters, whereby beneficial taxa such as *S. hominis* and *C. acnes* are lost throughout the pathogenesis of the disease.

S. aureus is known to dominate the skin microbiota of patients with atopic dermatitis and exacerbates the disease through inflammation [7,41]. Accordingly, a recent study of patients with new onset BP disease found BP lesions are frequently colonized by toxic shock syndrome toxin-1 producing *S. aureus* compared to age- and sex-matched controls [49]. Furthermore, Messingham et al. [49] describe a high rate of colonization with *S. aureus* in the nares and at healthy skin sites in their study population compared to controls, and that BP lesions were over six times more likely to be colonized than the nares or healthy skin from the same patients. Interestingly, there was no significant association between disease severity (BPDAI scores) or BP180/BP230 antibody levels with the type of colonizing bacteria, i.e., *S. aureus* or coagulase-negative staphylococci. However, the authors further observed that antibiotic therapy eliminated *S. aureus* and improved clinical outcomes.

In contrast, our large-scale investigation, which includes 228 patients with BP disease, sampling swabs from two anatomically distinct BP lesions, five control sites from each patient as well as matched control samples, lends high statistical power for detecting possible significant differences between cohorts. Consequently, we find *S. aureus* relative abundance to significantly positively correlate with disease severity at perilesional and contralateral sites. Moreover, *S. aureus* positively associates with BP disease regardless of sampling site, and the mean relative abundance of *S. aureus* is increased in sites rarely affected by BP as well as in the perilesional and contralateral sites. Importantly, we also find that disease severity negatively correlates with the coagulase-negative taxon *S. hominis*, even at sites without lesions. In sum, we believe our findings bolster those described by Messingham et al. [49], and

collectively suggest that *S. aureus* is an important indicator of BP. The specific role of this microbe and its functional components, e.g., how it might drive blister formation, will require further exploration.

In addition to cutaneous micro-environmental differences, geographic locations of patients with BP should be considered, as there is significant global variation in microbial colonization, especially as it relates to disease susceptibility [50–53]. Population differences observed in the gut microbiota in patients with inflammatory bowel disease, for example, suggest a complex interplay between geography and gut diseases that are in part driven by microbial factors [53]. Therefore, a broad-scale sampling of patients with BP across regions with variable incidences could reveal population-specific characteristics that might affect disease predisposition. We found that BPDAI scores explained a portion of microbial taxon variation between study centers. We recognize that geography represents an assemblage of factors including diet, culture, ancestry, and environmental features. Our results suggest the need for a large, global study to disentangle the relative importance of these features on the assembly of the skin microbiota, especially as it pertains to disease onset in AIBD.

Conclusions

In summary, our study suggests that skin microbiota may play an important role in the emergence of BP skin lesions, perhaps via the loss of beneficial taxa such as *S. hominis* and/or via the colonization of inflammation producing taxa such as *S. aureus*. Given the clear discriminatory power provided by differences in a few key indicator taxa, their relative proportions have the potential to provide critical information for assessing blistering risk as well as treatment outcomes. Future research may focus on functional analysis of host-microbe and microbe-microbe interactions as a means to identify novel treatment approaches for BP.

Funding sources

This work was supported by the German Research Foundation (DFG), through Clinical Research Unit 303, project number 269234613, subproject P2, jointly awarded to Prof. Dr. John F. Baines and Prof. Dr. Dr. Enno Schmidt, and the DFG Cluster of Excellence 2167 'Precision Medicine in Chronic Inflammation (PMI)' (grant no. EXC2167).

Data availability

Datasets related to this article can be found under BioProject accession number PRJNA715468 at https://www.ncbi.nlm.nih.gov/bioproject/, an open-source online data repository hosted at the NCBI SRA BioProject database. Reviewer link: https://dataview.ncbi.nlm.nih.gov/object/PRJNA715468?reviewer=tpob0l00b03qg88ju4lqnp7dou.

Compliance with ethics requirement

This research was approved by the University of Lübeck ethics committee (15–051, 18–046), as well as the respective committees of the study centers, following the Declaration of Helsinki. Written, informed consent was obtained from each participant.

CRediT authorship contribution statement

Meriem Belheouane: Data curation, Formal analysis, Investigation, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Britt M. Hermes:** Data curation, Formal analysis, Investigation, Software, Visualization, Writing – original draft, Writing – review & editing. **Nina Van Beek:** Project administration, Resources, Writing – review & editing. **Sandrine Benoit:** Resources. **Philippe Bernard:** Resources, Writing – review & editing. **Kossara Drenovska:** Resources, Writing – review & editing. **Sascha Gerdes:** Resources, Writing – review & editing. **Regine Gläser:** Resources, Writing – review & editing. **Matthias Goebeler:** Resources, Writing – review & editing. **Claudia Günther:** Resources, Writing – review & editing. **Anabelle von Georg:** Resources, Writing – review & editing. **Christoph M. Hammers:** Project administration, Resources, Writing – review & editing. **Maike M. Holtsche:** Project administration, Resources, Writing – review & editing. **Bernhard Homey:** Resources, Writing – review & editing. **Orsolya N. Horváth:** Writing – review & editing. **Franziska Hübner:** Resources, Writing – review & editing. **Beke Linnemann:** Resources, Writing – review & editing. **Pascal Joly:** Resources, Writing – review & editing. **Dalma Márton:** Writing – review & editing. **Aikaterini Patsatsi:** Resources, Writing – review & editing. **Claudia Pföhler:** Resources. **Miklós Sárdy:** Resources, Writing – review & editing. **Laura Huilaja:** Resources, Writing – review & editing. **Snejina Vassileva:** Resources, Writing – review & editing. **Detlef Zillikens:** Writing – review & editing. **Saleh Ibrahim:** Conceptualization, Writing – review & editing. **Christian D. Sadik:** Conceptualization, Funding acquisition, Methodology, Resources, Writing – review & editing. **Enno Schmidt:** Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Writing – review & editing. **John F. Baines:** Funding acquisition, Methodology, Project administration, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to thank Jan Schubert, Katja Cloppenborg-Schmidt, and Olga Eitel for their excellent technical assistance. We are grateful to Sarah Gaugel and Stephanie Freyher, Lübeck, for technical assistance with sample storage. We are indebted to Ana Luiza Lima, Kaan Yilmaz, and Onur Dikmen, Lübeck, for assistance with sample storage and communication with study centers during various phases of the study.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jare.2022.03.019.

References

[1] Joly P, Baricault S, Sparsa A, Bernard P, Bédane C, Duvert-Lehembre S, et al. Incidence and Mortality of Bullous Pemphigoid in France. J Investigative Dermatol 2012;132:1998–2004. doi: https://doi.org/10.1038/jid.2012.35.
[2] Schmidt E, Zillikens D. Pemphigoid diseases. The Lancet 2013;381:320–32. doi: https://doi.org/10.1016/S0140-6736(12)61140-4.
[3] Beek N, Weidinger A, Schneider SW, Kleinheinz A, Gläser R, Holtsche MM, et al. Incidence of pemphigoid diseases in Northern Germany in 2016 – first data from the Schleswig-Holstein Registry of Autoimmune Bullous Diseases. J Eur Acad Dermatol Venereol 2021;35(5):1197–202.
[4] Stevens NE, Cowin AJ, Kopecki Z. Skin Barrier and Autoimmunity—Mechanisms and Novel Therapeutic Approaches for Autoimmune Blistering Diseases of the Skin. Front Immunol 2019;10. doi: https://doi.org/10.3389/fimmu.2019.01089.

[5] Amber KT, Murrell DF, Schmidt E, Joly P, Borradori L. Autoimmune Subepidermal Bullous Diseases of the Skin and Mucosae: Clinical Features, Diagnosis, and Management. Clin Rev Allergy Immunol 2018;54:26–51. doi: https://doi.org/10.1007/s12016-017-8633-4.
[6] Huang S, Haiminen N, Carrieri A-P, Hu R, Jiang L, Parida L, et al. Human Skin, Oral, and Gut Microbiomes Predict Chronological Age. mSystems 2020;5(1).
[7] Kong HH, Oh J, Deming C, Conlan S, Grice EA, Beatson MA, et al. Temporal shifts in the skin microbiome associated with disease flares and treatment in children with atopic dermatitis. Genome Res 2012;22:850–9. doi: https://doi.org/10.1101/gr.131029.111.
[8] Quan C, Chen X-Y, Li X, Xue F, Chen L-H, Liu N, et al. Psoriatic lesions are characterized by higher bacterial load and imbalance between Cutibacterium and Corynebacterium. J Am Acad Dermatol 2020;82:955–61. doi: https://doi.org/10.1016/j.jaad.2019.06.024.
[9] Yerushalmi M, Elalouf O, Anderson M, Chandran V. The skin microbiome in psoriatic disease: A systematic review and critical appraisal | Elsevier Enhanced Reader. J Transl Autoimmunity 2019;2:100009.
[10] Salava A, Lauerma A. Role of the skin microbiome in atopic dermatitis. Clin Transl Allergy 2014;4:33. doi: https://doi.org/10.1186/2045-7022-4-33.
[11] Nakamizo S, Egawa G, Honda T, Nakajima S, Belkaid Y, Kabashima K. Commensal bacteria and cutaneous immunity. Semin Immunopathol 2015;37:73–80. doi: https://doi.org/10.1007/s00281-014-0452-6.
[12] Gao Z, Tseng C-H, Strober BE, Pei Z, Blaser MJ, Ahmed N. Substantial Alterations of the Cutaneous Bacterial Biota in Psoriatic Lesions. PLoS ONE 2008;3(7):e2719.
[13] O'Neill AM, Gallo RL. Host-microbiome interactions and recent progress into understanding the biology of acne vulgaris. Microbiome 2018;6:177. doi: https://doi.org/10.1186/s40168-018-0558-5.
[14] Schwarm C, Gola D, Holtsche MM, Dieterich A, Bhandari A, Freitag M, et al. German AIBD Study Group, Identification of two novel bullous pemphigoid-associated alleles, HLA-DQA1*05:05 and -DRB1*07:01, in Germans. Orphanet J Rare Dis 2021;16:228. doi: https://doi.org/10.1186/s13023-021-01863-9.
[15] Liu S-D, Chen W-T, Chi C-C. Association Between Medication Use and Bullous Pemphigoid: A Systematic Review and Meta-analysis. JAMA Dermatol 2020;156:891–900. doi: https://doi.org/10.1001/jamadermatol.2020.1587.
[16] Försti A-K, Huilaja L, Schmidt E, Tasanen K. Neurological and psychiatric associations in bullous pemphigoid-more than skin deep? Exp Dermatol 2017;26:1228–34. doi: https://doi.org/10.1111/exd.13401.
[17] Srinivas G, Möller S, Wang J, Künzel S, Zillikens D, Baines JF, et al. Genome-wide mapping of gene–microbiota interactions in susceptibility to autoimmune skin blistering. Nat Commun 2013;4:1–7. doi: https://doi.org/10.1038/ncomms3462.
[18] Miodovnik M, Künstner A, Langan EA, Zillikens D, Gläser R, Sprecher E, et al. A distinct cutaneous microbiota profile in autoimmune bullous disease patients. Exp Dermatol 2017. doi: https://doi.org/10.1111/exd.13357.
[19] Feliciani C, Joly P, Jonkman MF, Zambruno G, Zillikens D, Ioannides D, et al. Management of bullous pemphigoid: the European Dermatology Forum consensus in collaboration with the European Academy of Dermatology and Venereology. Br J Dermatol 2015;172:867–77. doi: https://doi.org/10.1111/bjd.13717.
[20] Schmidt E, Goebeler M, Hertl M, Sárdy M, Sitaru C, Eming R, et al. S2k guideline for the diagnosis of pemphigus vulgaris/foliaceus and bullous pemphigoid. J Dtsch Dermatol Ges 2015;13:713–27. doi: https://doi.org/10.1111/ddg.12612.
[21] E. Schmidt, R. Groves, Immunobullous diseases, in: Rook's Textbook of Dermatology, 9th ed., Wiley-Blackwell, Chichester, UK, 2016: p. 50.01-56.
[22] The Human Microbiome Project Consortium, Structure, function and diversity of the healthy human microbiome, Nature. 486 (2012) 207–214. https://doi.org/10.1038/nature11234.
[23] Karstens L, Asquith M, Davin S, Fair D, Gregory WT, Wolfe AJ, et al. Controlling for Contaminants in Low-Biomass 16S rRNA Gene Sequencing Experiments. mSystems 2019;4.
[24] Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform. Appl Environ Microbiol 2013;79:5112–20. doi: https://doi.org/10.1128/AEM.01043-13.
[25] Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. Microbiome 2018;6:226. doi: https://doi.org/10.1186/s40168-018-0605-2.
[26] Kong HH, Segre JA. The Molecular Revolution in Cutaneous Biology: Investigating the Skin Microbiome. J Investigative Dermatol 2017;137:e119–22. doi: https://doi.org/10.1016/j.jid.2016.07.045.
[27] Glassing A, Dowd SE, Galandiuk S, Davis B, Chiodini RJ. Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. Gut Pathog 2016;8:24. doi: https://doi.org/10.1186/s13099-016-0103-7.
[28] Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol 2014;12:87. doi: https://doi.org/10.1186/s12915-014-0087-z.
[29] Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods 2016;13:581–3. doi: https://doi.org/10.1038/nmeth.3869.
[30] Belheouane M, Vallier M, Čepić A, Chung CJ, Ibrahim S, Baines JF. Assessing similarities and disparities in the skin microbiota between wild and laboratory populations of house mice. ISME J 2020;14(10):2367–80.
[31] McMurdie PJ, Holmes S, Watson M. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS ONE 2013;8(4):e61217.
[32] Weyrich LS, Farrer AG, Eisenhofer R, Arriola LA, Young J, Selway CA, et al. Laboratory contamination over time during low-biomass sample analysis. Mol Ecol Res 2019;19:982–96. doi: https://doi.org/10.1111/1755-0998.13011.
[33] Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 2013;41:D590–6. doi: https://doi.org/10.1093/nar/gks1219.
[34] Wang Q, Garrity GM, Tiedje JM, Cole JR. Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. Appl Environ Microbiol 2007;73:5261–7. doi: https://doi.org/10.1128/AEM.00062-07.
[35] Cole JR, Chai B, Farris RJ, Wang Q, Kulam SA, McGarrell DM, et al. The Ribosomal Database Project (RDP-II): sequences and tools for high-throughput rRNA analysis. Nucleic Acids Res 2005;33:D294–6. doi: https://doi.org/10.1093/nar/gki038.
[36] Murrell DF, Daniel BS, Joly P, Borradori L, Amagai M, Hashimoto T, et al. Definitions and outcome measures for bullous pemphigoid: Recommendations by an international panel of experts. J Am Acad Dermatol 2012;66(3):479–85.
[37] J. Oksanen, F.G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P.R. Minchin, R.B. O'Hara, G.L. Simpson, P. Solymos, M.H.H. Stevens, E. Szoecs, H. Wagner, Community Ecology Package, 2005. http://sortie-admin.readyhosting.com/lme/R%20Packages/vegan.pdf (accessed March 10, 2021).
[38] Cáceres MD, Legendre P. Associations between species and groups of sites: indices and statistical inference. Ecology 2009;90:3566–74. doi: https://doi.org/10.1890/08-1823.1.
[39] Liaw A, Wiener M. Classification and Regression by randomForest. Newsl R Project 2002;2:5.
[40] Scholz CFP, Kilian M. The natural history of cutaneous propionibacteria, and reclassification of selected species within the genus Propionibacterium to the proposed novel genera Acidipropionibacterium gen. nov., Cutibacterium gen. nov. and Pseudopropionibacterium gen. nov. Int J Syst Evol Microbiol 2016;66:4422–32. doi: https://doi.org/10.1099/ijsem.0.001367.
[41] Nakatsuji T, Chen TH, Narala S, Chun KA, Two AM, Yun T, et al. Antimicrobials from human skin commensal bacteria protect against Staphylococcus aureus and are deficient in atopic dermatitis. Sci Transl Med 2017;9. doi: https://doi.org/10.1126/scitranslmed.aah4680.
[42] Clausen M-L, Agner T, Lilje B, Edslev SM, Johannesen TB, Andersen PS. Association of Disease Severity With Skin Microbiome and Filaggrin Gene Mutations in Adult Atopic Dermatitis. JAMA Dermatol 2018;154:293–300. doi: https://doi.org/10.1001/jamadermatol.2017.5440.
[43] Oh J, Byrd AL, Deming C, Conlan S, Kong HH, Segre JA. Biogeography and individuality shape function in the human skin metagenome. Nature 2014;514(7520):59–64.
[44] McLaughlin J, Watterson S, Layton AM, Bjourson AJ, Barnard E, McDowell A. Propionibacterium acnes and Acne Vulgaris: New Insights from the Integration of Population Genetic, Multi-Omic, Biochemical and Host-Microbe Studies. Microorganisms 2019;7(5):128.
[45] Gribbon EM, Cunliffe WJ, Holland KT. Interaction of Propionibacterium acnes with skin lipids in vitro. Microbiology 1993;139:1745–51. doi: https://doi.org/10.1099/00221287-139-8-1745.
[46] Nakamura K, O'Neill AM, Williams MR, Cau L, Nakatsuji T, Horswill AR, et al. Short chain fatty acids produced by Cutibacterium acnes inhibit biofilm formation by Staphylococcus epidermidis. Sci Rep 2020;10:21237. doi: https://doi.org/10.1038/s41598-020-77790-9.
[47] Shu M, Wang Y, Yu J, Kuo S, Coda A, Jiang Y, et al. Fermentation of Propionibacterium acnes, a Commensal Bacterium in the Human Skin Microbiome, as Skin Probiotics against Methicillin-Resistant Staphylococcus aureus. PLoS ONE 2013;8(2):e55380.
[48] Nakatsuji T, Hata TR, Tong Y, Cheng JY, Shafiq F, Butcher AM, et al. Development of a human skin commensal microbe for bacteriotherapy of atopic dermatitis and use in a phase 1 randomized clinical trial. Nat Med 2021;27(4):700–9.
[49] Messingham KN, Cahill MP, Kilgore SH, Munjal A, Schlievert PM, Fairley JA. TSST-1+Staphylococcus aureus in Bullous pemphigoid. J Invest Dermatol 2022;142(4):1032–1039.e6.
[50] Chen YE, Tsao H. The skin microbiome: current perspectives and future challenges. J Am Acad Dermatol 2013;69:143–155.e3. doi: https://doi.org/10.1016/j.jaad.2013.01.016.
[51] Hospodsky D, Pickering AJ, Julian TR, Miller D, Gorthala S, Boehm AB, et al. Hand bacterial communities vary across two different human populations. Microbiology 2014;160:1144–52. doi: https://doi.org/10.1099/mic.0.075390-0.
[52] Blaser MJ, Dominguez-Bello MG, Contreras M, Magris M, Hidalgo G, Estrada I, et al. Distinct cutaneous bacterial assemblages in a sampling of South American Amerindians and US residents. ISME J 2013;7:85–95. doi: https://doi.org/10.1038/ismej.2012.81.
[53] Rehman A, Rausch P, Wang J, Skieceviciene J, Kiudelis G, Bhagalia K, et al. Geographical patterns of the standing and active human gut microbiome in health and IBD. Gut 2016;65:238–48. doi: https://doi.org/10.1136/gutjnl-2014-308341.

Chapter 2

Skin microbiota analysis in patients with anorexia nervosa and healthy-weight controls reveals microbial indicators of healthy weight and associations with the antimicrobial peptide psoriasin.

Hermes, B. M., Rademacher, F., Chung, C. J., Tiegs, G., Bendix, M. C., de Zwaan, M., Harder, J., Baines, J. F. Skin microbiota analysis in patients with anorexia nervosa and healthy-weight controls reveals microbial indicators of healthy weight and associations with the antimicrobial peptide psoriasin. ***Scientific Reports.*** Submitted; under review. 2022.

Skin microbiota composition and correlation to antimicrobial peptides in patients with anorexia nervosa and healthy-weight controls

Hermes, Britt M.[1,2,3*], Rademacher, Franziska[4], Chung, Cecilia[1], Tiegs, Gisa[5], Bendix, Marie-Christin[6], de Zwaan, Martina[6], Harder, Jürgen[4*+], Baines, John F.[1,2*+]

Author affiliations
[1] Max Planck Institute for Evolutionary Biology, Plön, Germany
[2] Section of Evolutionary Medicine, Institute for Experimental Medicine, Kiel University, Kiel, Germany
[3] Lübeck Institute of Experimental Dermatology, University of Lübeck, Lübeck, Germany
[4] Department of Dermatology, University Hospital Schleswig-Holstein, Kiel, Germany
[5] Institute of Experimental Immunology and Hepatology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany.
[6] Department of Psychosomatic Medicine and Psychotherapy, Hannover Medical School, Germany
[+] These authors jointly supervised this work: Harder, Jürgen and Baines, John F.
[*] Corresponding authors: hermes@evolbio.mpg.de; baines@evolbio.mpg.de; jharder@dermatology.uni-kiel.de

Abbreviations: anorexia nervosa (AN), antimicrobial peptide (AMP), body mass index (BMI), amplicon-sequence variants (ASV)

Abstract
Anorexia nervosa (AN), a psychiatric condition defined by low body weight for age and height, is associated with numerous dermatological conditions. Yet, clinical observations report that patients with AN do not suffer from infectious skin diseases like those associated with primary malnutrition. Cell-mediated immunity appears to be amplified in AN, however, this pro-inflammatory state does not sufficiently explain the lower incidence of infections. Antimicrobial peptides (AMPs) are important components of the innate immune system protecting from pathogens and shaping the microbiota. In *Drosophila melanogaster* starvation precedes increased AMP gene expression. Here, we analyzed skin microbiota in patients with AN and age-matched, healthy-weight controls and investigated the influence of weight gain on microbial community structure. We then correlated features of the skin microbial community with psoriasin and RNase 7, two highly abundant AMPs in human skin, to clarify whether an association between AMPs and skin microbiota exists and whether such a relationship might contribute to the resistance to cutaneous infections observed in AN. We find significant statistical correlations between Shannon diversity and the highly abundant skin AMP psoriasin and bacterial load, respectively. Moreover, we reveal psoriasin significantly associates with *Abiotrophia*, an indicator for the healthy-weight control group. Additionally, we observe a significant correlation between an individual's body mass index and *Lactobacillus*, a microbial indicator of health. Future investigation may help clarify physiological mechanisms that link nutritional intake with skin physiology.

Introduction

Anorexia nervosa (AN) is a psychiatric condition typically affecting females with an estimated lifetime prevalence between 0.5% to 2.0% and a peak age in onset between 13 and 18 years of age[1]. The hallmark feature of AN is low body weight for age and height, usually achieved via extreme caloric restriction. AN is complicated by malnutrition that can lead to life-threatening medical consequences as a result of multiple organ failure and immune system dysfunction[2,3].

Starvation, malnutrition, altered dietary patterns, and single-nutrient deficiencies can all cause impaired immune functioning that can lead to chronic inflammation and recurrent infections[4,5]. Indeed, malnourished children chiefly die from "common infections[4,6]." Obese individuals, who often have micronutrient deficiencies, experience more frequent and more severe infections[4]. Paradoxically, clinical observations of patients with AN report an absence of infections as well as delayed or reduced physiological responses to infection[7–11]. Moreover, AN associated dermatological changes include xerosis (dry skin), increased acne, slower wound healing, generalized pruritis, and seborrheic dermatitis, but an increased risk of skin infections has not been reported[12]. This is in striking contrast to an increased risk of skin infections associated with primary malnutrition typically seen in developing nations[13,14]. Cell-mediated immunity appears to be amplified in AN, however, this pro-inflammatory state does not sufficiently explain the lower incidence of infections[2,11,15,16].

Antimicrobial peptides (AMPs) are evolutionarily conserved effector molecules of the innate immune system with broad-spectrum antimicrobial activities[17]. Psoriasin and RNase 7 are the most abundant AMPs found on human skin that serve immunomodulatory roles in skin immunity through the induction of cytokines and chemokines[18,19]. In the chronic skin inflammatory disease psoriasis, keratinocytes proliferate in response to inflammatory cytokines, which in turn increases the synthesis of AMPs, including psoriasin, and contributes to the recruitment of T cell subsets and other immune effector cells into the skin[20,21]. RNase 7 is induced by proinflammatory cytokines and a wide spectrum of potential pathogenic microorganisms such as *Staphylococcus aureus* and *Candida albicans*[19,22]. Similar to psoriasin, RNase 7 is upregulated in psoriasis and atopic dermatitis[23,24]. It is thus intriguing that the expression of AMP genes are also induced by starvation in *Drosophila melanogaster* (common fruit fly) in the absence of infection and independent of the pathogen-response pathway[25–27]. It is possible that this mechanism evolved to ensure innate immune activity during periods of energy deprivation.

Previously, to evaluate whether weight status may also affect AMP expression in human skin, we analyzed the concentrations of the AMPs psoriasin and RNase 7 on the skin surface of patients with AN before and after weight gain. Surprisingly, we found AMP concentrations did not decrease with weight gain, but rather an association of weight gain with increasing AMP concentrations was observed[28]. While a link between AN and skin immune function has yet to be elucidated, we hypothesize here that changes in the skin microbial profile of patients with AN might contribute to the absence of skin infections observed in this population.

In this study, we conducted an analysis of skin microbiota based on 16S rRNA gene amplicon sequencing in female patients with AN before and after undergoing an in-patient treatment program to gain weight and compared to age-matched healthy-weight controls. To test for possible relationships between AMP concentrations, bacterial load, or body mass index (BMI) and skin microbiota, and to gain insight into whether such relationships might contribute to the resistance to dermatological infections observed in AN, we analyzed skin microbial profiles in conjunction with these measures. We observe increasing levels of bacterial load with weight gain in patients with AN, which is significant at the inner elbow sampling location. We reveal significant correlations between psoriasin concentrations and the indicator taxon *Abiotrophia* for the healthy-weight control group. Further, we find Shannon diversity significantly negatively

correlates with psoriasin concentrations as well as total bacterial load. Finally, we observe a significant correlation between an individual's BMI and *Lactobacillus*, a significant microbial indicator of health.

Results

Study participants and skin sampling

Thirty-three females diagnosed with AN receiving inpatient medical care and thirty-three healthy-weight age-matched female control subjects from Germany were recruited for this study (see Methods for inclusion and exclusion details). One patient with AN withdrew from the study prior to the second sampling point. Patient metadata analyzed in this study are summarized in Table 1; complete study metadata are provided online in Supplementary Table S1. In patients with AN, the mean BMI was 12.56 kg/m^2 (SD 1.7) before weight gain and 14.54 kg/m^2 (SD 1.7) after weight gain, with a mean weight gain of 5.7 kg (SD 1.5) corresponding to an increase in BMI of 2.0 (SD 0.5) points. The mean BMI of the healthy-weight control group was 22.10 kg/m^2 (1.73). All patients with AN had been diagnosed with a severe and life-threatening stage of AN, according to DSM-5 criteria, as represented by a BMI of 15.0 kg/m^2 or less.

Table 1. Summarized metadata

	AN (n=33)	HC subjects (n=33)
Age at time of first measurement		
Mean (SD)	25.8 yrs (9.9)	26.0 yrs (9.9)
Median	23	22.5
Range	17-54	16-57
Weight (kg) first measurement, before weight gain		
Mean (SD)	35.4 (5.2)	61.55 (6.18)
Median	36.9	61
Range	23.4-45.5	49-80
BMI first measurement, before weight gain		
Mean (SD)	12.56 (1.70)	22.10 (1.73)
Median	12.87	22.28
Range	9.22-14.8	18.69-25.84
	AN (n=32)	**HC subjects (n=33)**
Weight (kg) second measurement, after weight gain		
Mean (SD)	40.3 (5.2)	-
Median	42.4	-
Range	29-47	-
BMI second measurement, after weight gain		
Mean (SD)	14.54 (1.72)	-
Median	14.81	-
Range	10.73-17.55	-

AN = patients with anorexia nervosa
HC = healthy weight controls

To obtain skin bacterial profiles and bacterial load estimates, we extracted DNA from material derived from a skin-rinsing protocol that concurrently collected the antimicrobial peptides psoriasin and RNase 7 from three standardized body sites representing sebaceous, moist, and dry cutaneous zones (Fig. 1). Sampling sites were 1.77 cm^2 in size and included the forehead

(sebaceous), the antecubital fossa (referred to as inner elbow in this study; moist), and the ventral side of the lower forearm (dry). Patients were positioned accordingly to facilitate sampling procedures, e.g., were placed supine for forehead sampling.

For the subsequent analyses, we defined three study subject groupings: i) healthy-weight controls (HC) defined by a BMI ranging between 18.5 and 25.0 kg/m^2, ii) patients with AN prior to gain weight (hereafter referred to as AN before weight gain), and iii) patients with AN after undergoing an inpatient protocol to gain weight and after having gained at least 2 kg of body cell mass (hereafter referred to as AN after weight gain). Accordingly, we analyzed these three groups according to total bacterial load derived from digital droplet PCR (ddPCR), the relative abundance of major taxa and diversity patterns identified in the 16S rRNA gene analysis, and our previously published concentrations of psoriasin and RNase 7[28]. A summary of the mean and median concentrations of psoriasin and RNase 7 is provided in Supplementary Table S2.

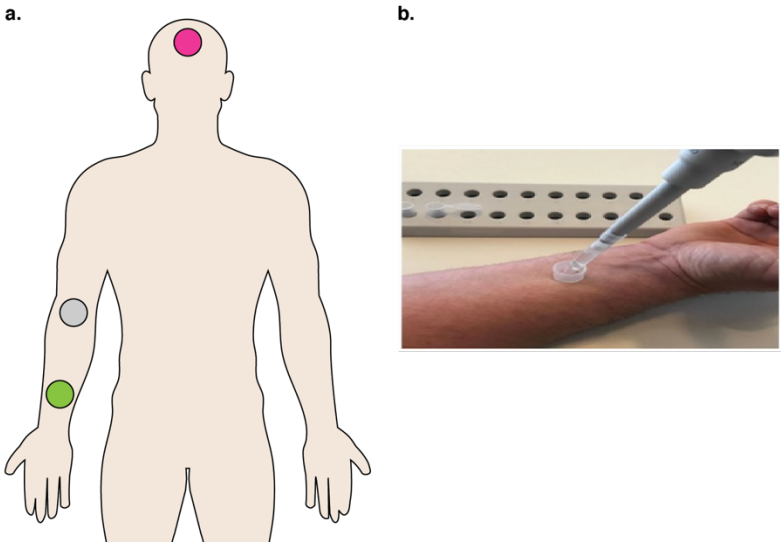


Fig. 1. Overview of sampling procedure used in this study. **a**. Standardized sampling locations for healthy-weight controls and patients with AN included the forehead, a sebaceous zone (pink), the antecubital fossa/ inner elbow, a moist zone (grey), and the ventral side of the lower arm, a dry zone (green). Patients were positioned accordingly to facilitate sampling procedures, e.g., supine for forehead sampling. Illustration by B. Hermes, 2021. **b.** Sampling procedure of the distal (lower) forearm. Image photographed by Bendix *et al*., 2020.

Bacterial load

Due to the low microbial biomass of the skin environment and the associated technical challenges[29], and the reasonable expectation that AMPs and/or AN disease status could influence bacterial load, we initially measured the total bacterial load of each sample using ddPCR to obtain a precise quantification of target DNA copies, as described by Sze and colleagues[30] (see Methods) (Supplementary Table S3). ddPCR is a method whereby a sample is fractionated into tens of thousands of individual droplets using a water-oil emulsion; PCR is then carried out within each droplet thereby providing reliable, absolute quantification of the target molecule, reducing PCR bias, and increasing signal-to-noise ratios, especially in low biomass samples such as skin[30–35]. We assessed the distribution of bacterial load between groups at individual sampling locations (Fig. 2). We observe an overall trend of increasing bacterial load with weight gain in patients with AN. However, we largely find that differences in bacterial load between groups are not significant, except for at the inner elbow (antecubital fossa). Here, differences in load between HC and patients with AN after weight gain reach statistical significance but differences are not significant between patients with AN before and after weight gain (Wilcoxon; $p = 0.009$;

$p = 0.053$, respectively; Fig. 2). We additionally find that bacterial load significantly correlates with psoriasin concentrations at the forehead (Spearman; $r_s = 0.28$, $p = 0.02$), but not at the inner elbow or lower arm (Supplementary Table S3). Bacterial load did not significantly correlate with RNase 7 concentrations or BMI.

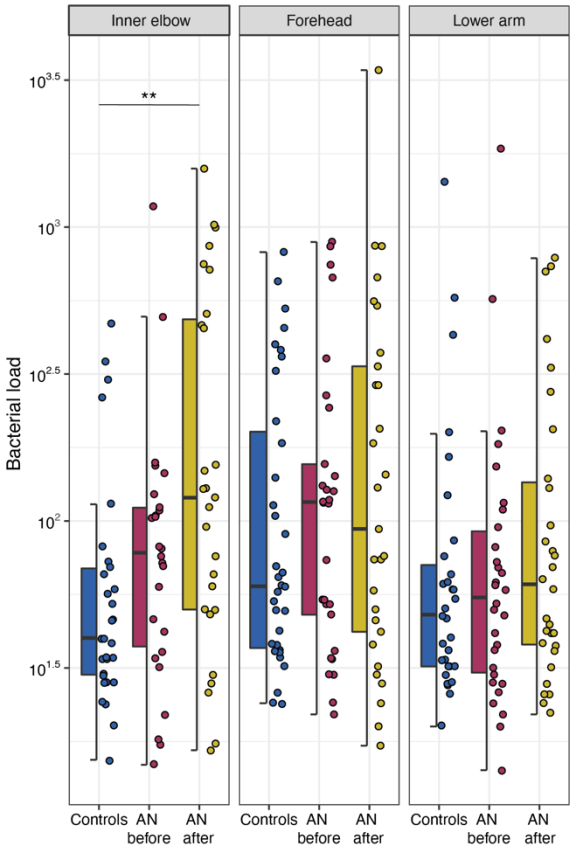


Figure 2. Box plots of total bacterial load, as measured by ddPCR, for study groups at individual sampling locations. AN = anorexia nervosa. Wilcoxon test (see Methods); *p*-values: * <0.05; ** < 0.01; *** < 0.001. *p*-values were adjusted for multiple testing according to Benjamini and Hochberg (1995). Line indicates the median concentration; box shows the interquartile range (IQR), and the whiskers are 1.5x IQR. Blue represents healthy-weight controls, red represents patients with AN before weight gain, and gold represents patients with AN after weight gain.

Next, we used ddPCR measurements to aid the assessment of potential contamination (see Methods for a detailed description). Briefly, total bacterial load was used as a proxy for input bacterial DNA concentrations for the "frequency" method within the R package "Decontam" (version 1.8.0; see Methods)[36,37]. To verify ASVs classified as contaminants (n = 154), we visualized five randomly selected contaminants in frequency plots to examine the distribution of the ASV with respect to bacterial loads. We find that the contaminant ASVs follow an expected pattern in which frequency is inversely proportional to bacterial load, as contaminating DNA will account for a larger fraction of this load in samples with low biomass[36]. We subsequently utilized the "prevalence" method within "Decontam" (version 1.8.0), in which the prevalence (presence /absence) of ASVs in samples is compared to the that in negative controls, to identify additional contaminants[36]. An additional 70 ASVs were identified as contaminants and removed from the dataset. Finally, following the recommendations of Weyrich *et al.*[38], any ASV belonging to families *Halomonadaceae* (n = 0) or *Shewanellaceae* (n = 14) were removed. In total, we analyzed more than 400,000 sequences, with a normalized coverage of 1,000 sequences per sample (see Methods).

Overview of skin microbiota in patients with AN and healthy-weight controls

We first analyzed community composition at the phylum and genus levels. The dominant phyla include Firmicutes, Actinobacteria, Proteobacteria, Bacteroidetes, and the dominant genera include *Staphylococcus*, *Streptococcus, Propionibacterium*, *Corynebacterium*, *Anaerococcus,* and *Lactobacillus*, whose relative proportions are shown in Fig. 3. Comparisons of relative abundances between groups at the phylum level revealed significant differences in Proteobacteria between patients with AN before and after weight gain and between HC and patients with AN before weight gain (Wilcoxon; $p = 0.005$, $p = 0.014$, respectively). Additionally, we find significant differences in Firmicutes abundance between HC and patients with AN before weight gain (Wilcoxon; $p = 0.003$). At the genus-level, there are significant differences in the relative abundance of *Lactobacillus* between HC and AN before weight gain, as well as between HC and AN after weight gain groups (Wilcoxon; $p = 4.23e-07$, 5.12e-08, respectively). Other significant differences between groups include *Staphylococcus* for AN before compared to HC (Wilcoxon; p = 0.021) and AN before compared to AN after (Wilcoxon; p = 0.005), unclassified *Neisseriaceae* for both AN before and AN after weight gain compared to HC (Wilcoxon; $p = 0.003$, $p = 0.03$, respectively) and unclassified *Streptophyta* for AN after weight gain compared to HC (Wilcoxon; $p = 0.05$). Supplementary Table S4 provides a summary of statistical analyses comparing the most abundant phyla and genera between groups.

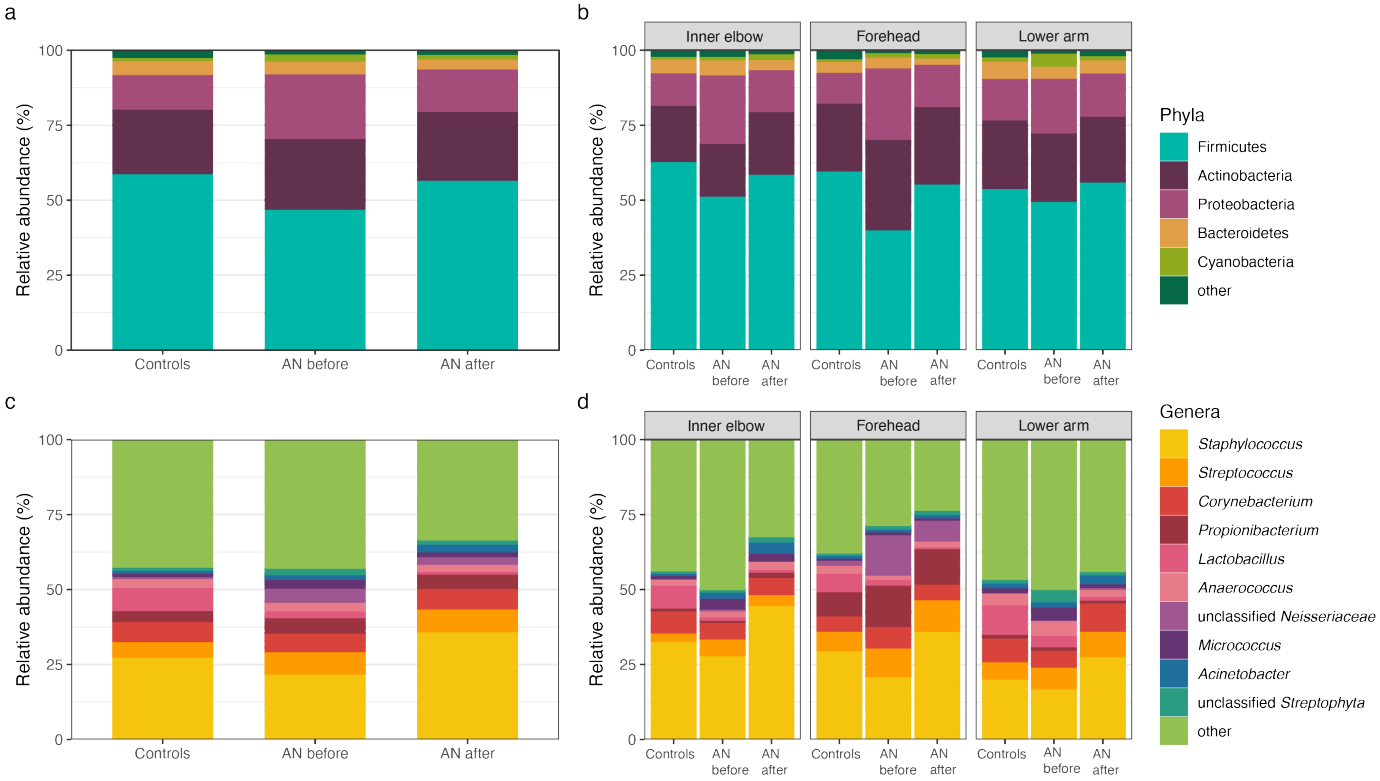


Fig. 3. Overview of dominant taxa at sampling sites. **a.** Bar plot of relative abundances for the most abundant phyla, and **b.** at sampling sites (inner elbow, forehead, and lower arm). **c.** Bar plot of relative abundances for the most abundant genera, and **d.** at sampling sites (elbow, forehead, and lower arm). AN = anorexia nervosa; Controls = healthy-weight controls; AN before = patients with AN before weight gain treatment; AN after = patients with AN after weight gain treatment

Diversity indices

Next, we assessed alpha diversity at the amplicon sequence variant (ASV)-level to investigate potential effects of AN on skin microbiota. Shannon diversity measures both the richness (number of different species) and evenness (how the species are distributed relative to one another) of the bacterial community, whereas the Chao1 index reflects expected species richness.

Surprisingly, we show alpha diversity tends to decrease in patients with AN after weight gain therapy and find significant differences in both community richness and evenness in these patients compared to HC. Specifically, we find a significant difference in Shannon diversity at the forehead between HC and AN after weight gain (Wilcoxon; $p = 0.02$; Fig. 4a). We also find a significant difference in Shannon diversity at the lower forearm between HC and AN after weight gain (Wilcoxon; $p = 0.005$; Fig. 4a). For Chao1 diversity, we find a significant difference at the forehead between AN before weight gain and HC and between AN after weight gain and HC (Wilcoxon; $p = 0.02$; $p = 0.007$, respectively; Fig. 4b). As with Shannon diversity, there is a significant difference in Chao1 diversity at the lower forearm between HC and AN after weight gain (Wilcoxon; $p = 0.009$; Fig. 4b). Supplementary Table S5 provides summary statistics for group comparisons.

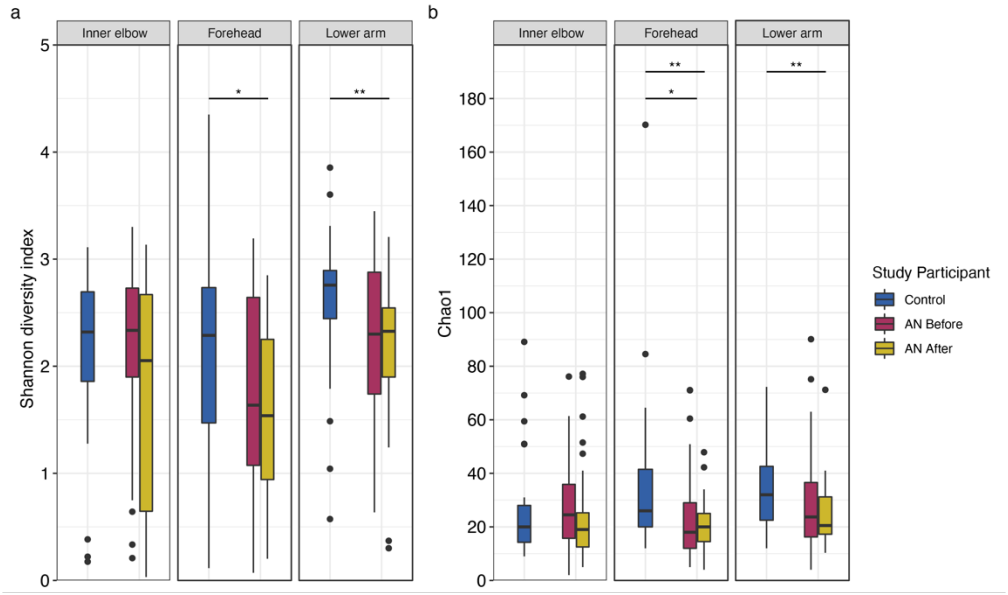


Fig. 4. Alpha diversity indices for healthy-weight controls and patients with anorexia nervosa (AN), by weight gain status (before and after), at each sampling location. **a.** Shannon diversity index. **b.** Chao1 diversity index. Wilcoxon test (see Methods**)**; *p*-values: * <0.05; ** < 0.01; *** < 0.001. *p*-values were adjusted for multiple testing according to Benjamini and Hochberg (1995). Line indicates the median concentration; box shows the interquartile range (IQR), and the whiskers are 1.5x IQR. Blue represents healthy-weight controls, red represents patients with AN before weight gain, and gold represents patients with AN after weight gain.

To further explore the trend of decreasing alpha diversity after weight gain in patients with AN, we calculated Spearman correlations for Shannon and Chao1 diversity measures with AMP concentrations, total bacterial load, and BMI (Supplementary Table S6). We find that Shannon diversity is significantly negatively correlated with psoriasin concentrations (Spearman; rho = -0.22, $p = 0.0003$; Fig 5a), but not within individual sampling sites (Fig. 5b). We find Shannon diversity also significantly negatively correlates with total bacterial load (Spearman; rho = -0.35, $p = 4.6e-09$; Fig. 5c), and moreover, this significant relationship is preserved at the inner elbow (Spearman; rho = -0.35, $p = 0.003$), the forehead (Spearman; rho = -0.33, $p = 0.003$), and the lower arm (Spearman; rho = -0.26, $p = 0.04$; Fig. 5d). Shannon diversity does not significantly covary with RNase 7 concentrations or BMI. Chao 1 index does not significantly correlate with AMP concentrations, total bacterial load, or BMI.

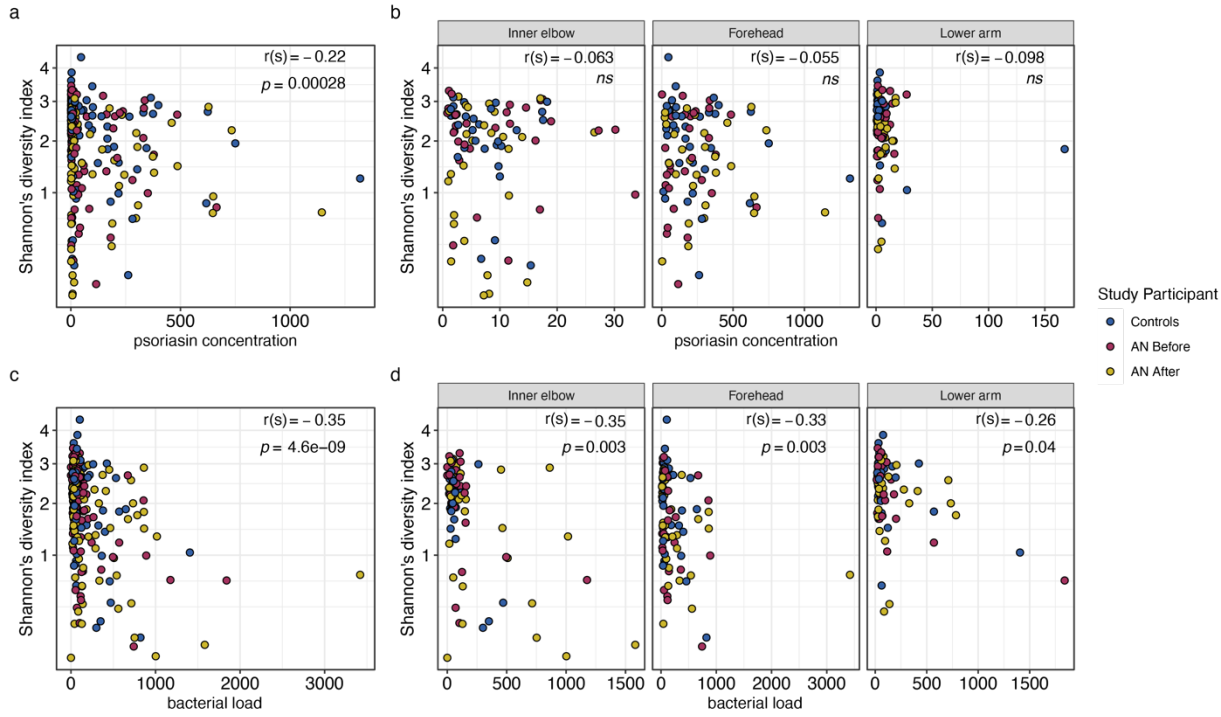


Fig. 5. Spearman correlations between Shannon diversity index and **a.** psoriasin concentrations, **b**. psoriasin concentrations at individual sampling locations, **c.** total bacterial load, and **d.** total bacterial load at individual sampling locations. r(s) = spearman's Rho. AN = anorexia nervosa. ns = not significant. Blue represents healthy-weight controls, red represents patients with AN before weight gain, and gold represents patients with AN after weight gain. *p*-values were adjusted for multiple testing according to Benjamini and Hochberg (1995).

To assess overall community compositional differences between groups, we next performed beta diversity analyses. We find no distinguishable separation of study groups based on the Bray-Curtis dissimilarity index (based on abundance) or with the Jaccard index (based on presence/absence), suggesting the similar microbial communities amongst the groups (Supplementary Fig. S1). A constrained analysis of principal coordinates of the Bray-Curtis distance ('*capscale*'[39]) with respect to treatment status (i.e., HC, AN before, and AN after) reveals significant differences between the study groups, but treatment status explains only about 1% of the variation between groups (anova.cca; p = 1e-04; Supplementary Fig. S2).

Indicator species

To reveal potentially important individual taxa, we conducted indicator species analyses ('*indicspecies*'[40]) at the ASV- and genus-level on a microbiota core defined by a prevalence threshold, whereby a taxon must be present in at least 5% of samples for inclusion in the analysis (see Methods).

At the genus-level, *Lactobacillus*, *Clostridium sensu stricto*, and *Abiotrophia* associate with the HC group (Table 2). Accordingly, there is a statistically significant difference in the relative abundance of *Lactobacillus* in HC compared to both the AN before and AN after weight gain (Wilcoxon; p = 4.23e-07, 5.12e-098, respectively; Fig.6a; Supplementary Table S4). These significant differences are maintained at individual sampling locations (Fig. 6b; Supplementary Table S4). Further, we identify unclassified *Neisseriaceae* as a significant indicator for both the AN before and AN after groups. Summary statistics for differences in indicator genera between groups at individual sampling locations within groups are presented in Supplementary Table S4.

Subsequently, because the "Decontam" procedure was performed on the level of ASVs, we conducted an additional screen to evaluate whether indicator genera could represent contaminants in our dataset, based on the expectation of negative correlations between bacterial load and contaminating taxa[36]. We thus accordingly calculated Spearman correlations between the relative abundance of indicator genera and bacterial loads. We find no significant correlations between total bacterial load and *Lactobacillus*, *Abiotrophia*, *Clostridium sensu stricto*, or unclassified *Neisseriaceae* (see Methods; Supplementary Table S7). However, our analysis finds *Jeotgalicoccus*, an additional indicator genus for the HC group, to negatively correlate with total bacterial load (Spearman; rs = -0.13, p = 0.03). This association is not significant at individual sampling locations. Nevertheless, following the logic that contaminant sequences are expected to negatively covary with bacterial loads, *Jeotgalicoccus* was excluded from additional analyses and not reported in Table 2.

Table 2. Indicators at genus- and ASV-level with RDP SeqMatch Results

	stat	p-value	adj. p-value	RDP SeqMatch result	S_ab score
Healthy controls					
Lactobacillus	0.706	2.00E-05	0.002	*Lactobacillus crispatus*	1.0
ASV_29	0.706	4.00E-05	0.005		
Abiotrophia	0.404	9.00E-04	0.03	*Abiotrophia defectiva*	1.0
ASV_160	0.404	0.001	0.03		
Clostridium sensu stricto	0.374	0.001	0.03	*Clostridium* spp.	1.0
ASV_744	0.374	0.001	0.03		
Patients with AN					
unclassified Neisseriaceae	0.502	0.002	0.05	NA	NA
ASV_13	0.502	0.003	0.05		

ASV = amplicon sequent variant; RDP = Ribosomal Database Project; AN = anorexia nervosa

Indicator species analysis applied using indicspecies (version 1.7.9) with "r.g." function and 99,999 permutations on a microbial core defined by ASVs classified to the genus-level that are present in at least 5% of all samples (see Methods). *p*-values were adjusted for multiple testing according to Benjamini and Hochberg (1995).

Representative 16S rRNA gene sequences were queried via Ribosomal Database Project SeqMatch (version 3).

At the ASV level, we find three significant indicators for the HC group, and one for both the patients with AN before and after weight gain (Table 2). To refine the taxonomic classification of indicator ASVs, we queried representative sequences using RDP SeqMatch (see Methods; Supplementary Table S8). Indicator ASV_29 is a close match to *Lactobacillus crispatus* (S_ab score = 1.0). *Lactobacilli* spp. are well-known human commensals, with previous studies reporting *Lactobacilli* spp. in the gut, vagina, mouth, on skin, and in breastmilk[41–43]. A query of indicator ASV_160 reveals a close match to *Abiotrophia defectiva* (S_ab score = 1.0). Previous studies identified *Abiotrophia* spp., from the family *Lactobacillales*, in the oral and upper respiratory flora[44]. *Clostridium sensu stricto* was identified in the human gut microbiome in the context of chronic disease[45] and was previously classified as a human-associated microbe with pathogenic capabilities[46].

To evaluate potential associations between the relative abundance of indicator genera with AMP concentrations and BMI, we calculated Spearman correlations (see Methods; Supplementary Table S7). We find that *Abiotrophia* significantly positively associates with psoriasin concentrations (Spearman; $r_s = 0.17$, $p = 0.004$; Supplementary Fig. S3). However, at individual sampling locations, i.e., elbow, forehead, and lower arm, these correlations are not significant (Supplementary Fig. S3). *Abiotrophia* does not significantly correlate with RNase 7 concentrations or BMI. We find that *Lactobacillus* does not correlate with psoriasin or RNase 7 concentrations. However, we find significant associations between BMI and *Lactobacillus* (Spearman; $r_s = 0.37$, $p = 5.4e\text{-}10$; Fig. 6c); further, this significant association is maintained at the inner elbow (Spearman; $r_s = 0.3$, $p = 0.005$), the forehead (Spearman; $r_s = 0.5$, $p = 1.65e\text{-}06$), and lower arm (Spearman; $r_s = 0.3$, $p = 0.001$; Fig. 6d). Since *Lactobacillus* is an abundant taxon, we selected it to conduct single linear regression modeling to assess whether BMI predicts the relative abundance of *Lactobacillus*. We found BMI to be a weak, but significant predictor of *Lactobacillus* relative abundance at the inner elbow ($R^2_{adj} = 0.08$; $p = 0.007$), the forehead ($R^2_{adj} = 0.05$; $p = 0.02$), and at the lower arm ($R^2_{adj} = 0.08$; $p = 0.007$; Fig. 6e).

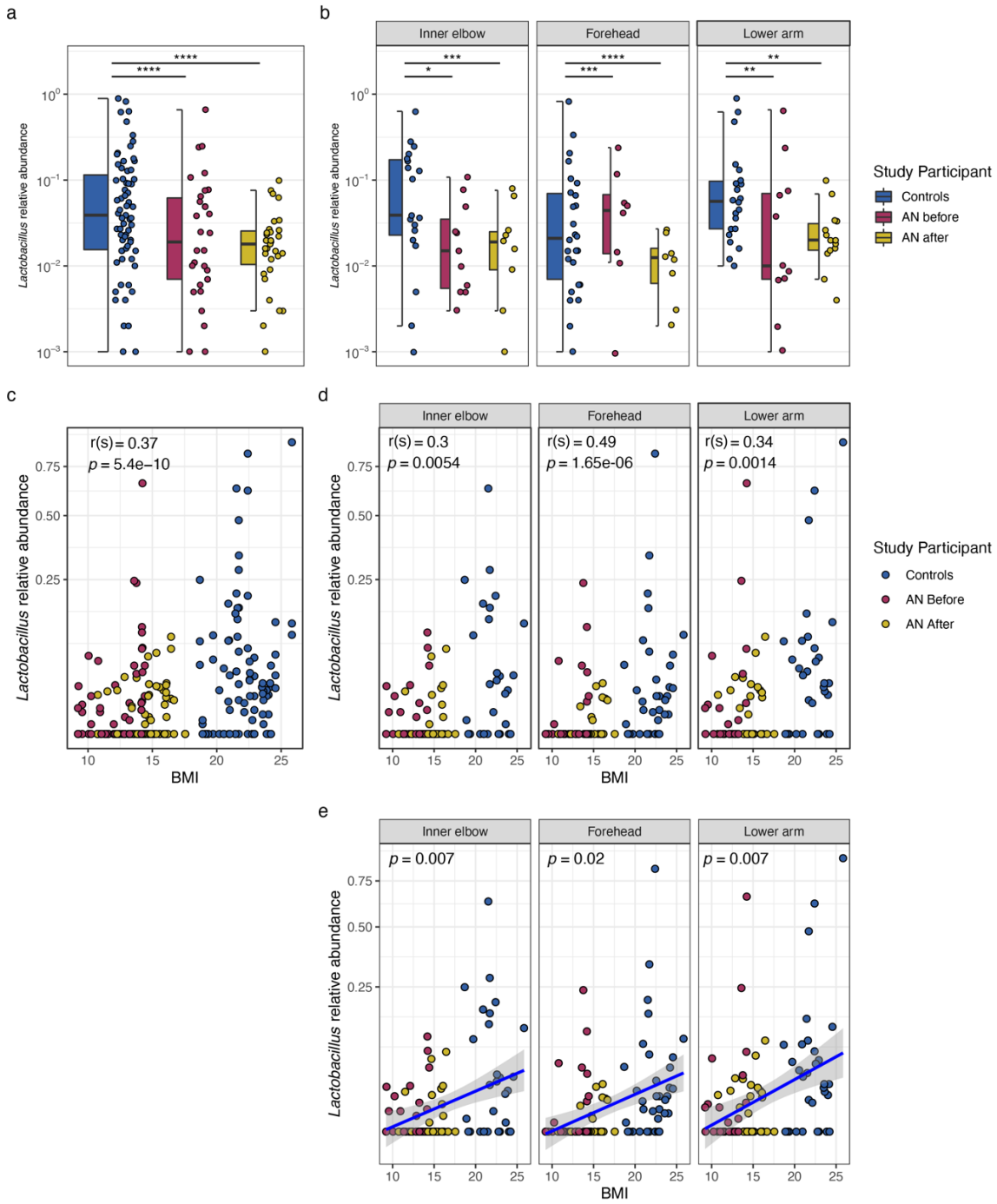


Figure 6. Box plots of relative abundances, Spearman correlations, and single linear regressions for *Lactobacillus*, an indicator genus. **a.** Box plot of *Lactobacillus* relative abundances for healthy-weight controls and patients with AN by weight gain arm (before and after) and, **b.** faceted by sampling location. Wilcoxon test (see Methods**)**; *p*-values: * <0.05; ** < 0.01; *** < 0.001. Line indicates the median concentration; box shows the interquartile range (IQR), and the whiskers are 1.5x IQR. **c.** Spearman correlation between indicator *Lactobacillus* and BMI, and at **d.** individual sampling locations **e.** Single linear regression model between BMI and relative abundance of *Lactobacillus* at the inner elbow (R^2_{adj} = 0.08; p = 0.007)*,* the forehead (R^2_{adj} = 0.05; p = 0.02), and at the lower arm (R^2_{adj} = 0.08; p = 0.007). r(s) = spearman's Rho. AN = anorexia nervosa. BMI = body mass index. Blue represents healthy-weight controls, red represents patients with AN before weight gain, and gold represents patients with AN after weight gain. *p*-values were adjusted for multiple testing according to Benjamini and Hochberg (1995).

Discussion

Our study is the first to characterize the skin microbiota in female patients with AN. We conducted a 16S rRNA gene-based analysis in patients with AN before and after weight gain and with age-matched, healthy-weight controls and then correlated these findings with the concentrations of two highly abundant skin AMPs, psoriasin and RNase 7, and with BMI.

Notably, we find that the concentration of the AMP psoriasin weakly, but significantly, correlates with the indicator genus *Abiotrophia* for HC. However, at individual sampling locations, this genus does not significantly correlate with psoriasin concentrations, possibly due to low frequencies and low relative abundances.

Recently, *Abiotrophia* was found to be positively associated with the severity of psoriasis, a mixed autoimmune and autoinflammatory skin disorder marked by elevated psoriasin concentrations[47]. In this regard, our findings that *Abiotrophia* significantly positively associates with psoriasin, an established biomarker for psoriasis, supports evidence that links *Abiotrophia* and psoriasis disease severity. In our study, the relative abundance of *Abiotrophia* represents less than 1% of the total abundance of skin microbiota. This is fitting, as our study population did not exhibit signs of inflammatory skin disease or psoriasis, and thus, we would not expect this taxon to be a dominant genus. Still, it is interesting that *Abiotrophia* is an indicator for HC, but not for patients with AN, where one might expect inflammation to occur alongside AN-associated skin changes. The role of cell-mediated immunity in AN is controversial, with several studies reporting an increase in T-cell proliferation and inflammatory cytokine production, including interleukin-1, interleukin-6, and tumor-necrosis factor, when compared to healthy controls or to innate immunity responses in primary malnutrition, where immune function is suppressed [2,11,15,16]. However, an earlier investigation conducted by Omodei and colleagues[48] found that immune cell populations and the cytokines they produce are reduced in AN, but display greater antioxidant, stress resistance, and anti-inflammatory profiles compared to controls. It is possible that the AN population included in our study exhibits an augmented anti-inflammatory profile, thereby clarifying the relatively reduced levels of microbial taxa associated with inflammatory skin disease, such as *Abiotrophia*, observed in these subjects.

We also find the Shannon diversity index, which reflects both species richness and species evenness, to significantly negatively associate with psoriasin concentrations and with total bacterial load. Furthermore, we show that both Shannon and Chao1 diversity decreases in patients with AN after weight gain compared to patients with AN before weight and compared to HC, and that there are significant differences in alpha diversity between HC and AN after weight gain. These results are interesting in the context of our previous work in which we found that AMP concentrations, psoriasin in particular, tended to increase in patients with AN after weight gain[28]. Previous surveys of the skin microbiome of patients with psoriasis report higher alpha diversity, but with lower stability, compared to healthy skin[49], while others report decreased taxonomic diversity in psoriatic skin compared to healthy skin[50]. Increased psoriasin expression is a well-established feature of psoriasis[51]. As such, it is possible that our findings capture a transitional shift in alpha diversity in patients with AN in response to rising psoriasin levels.

Additionally, we find that BMI significantly correlates with *Lactobacillus*, another indicator genus for HC. Since *Lactobacillus* represents a dominant genus in our study, we assessed the possibility that *Lactobacillus* is a contaminant. We visualized the distribution of *Lactobacillus* across bacterial loads obtained from ddPCR for the most abundant *Lactobacillus* ASVs in our dataset. We find that these *Lactobacillus* ASVs do not follow a pattern of contamination, whereby their frequency would be inversely proportional to input bacterial load. Rather, the frequency of *Lactobacillus* ASVs are independent of the input ddPCR load data. These findings are consistent with *Lactobacillus* representing a true biological signal.

The finding of *Lactobacillus,* and *L. crispatus* in particular, as an indicator of healthy-weight is congruent with previous studies demonstrating the potential role of *Lactobacilli* spp. as probiotics for improving skin health and barrier function[52,53], anti-aging effects[54], and balancing the gut microbial population, thereby preventing inflammatory disease and even cancer at different sites in the body, most likely through the production anti-inflammatory metabolites such as short chain fatty acids[55]. Moreover, this finding is supported by data showing that *Lactobacillus* colonizes healthy human skin[56–60], including the inner elbow[61], forehead[62], and scalp[63]. Interestingly, one study exploring the effects of age on the structure of the skin microbiome found *Lactobacillus* to be present on the skin of participants aged 20 to 30 years, but not on those aged 50 to 60 years[64]. This finding is particularly interesting given that the mean age of our study population is 25 years for patients with AN and 26 years for HC. However, another survey of the skin microbiome in relation to age and photodamage found increasing age is associated with an increase in *Lactobacillus*[60]. Li et al. (2020) hypothesize that *Lactobacillus* spp. may increase in response to skin damage (e.g., from UV radiation) that accumulates with age, which may reduce inflammation and improve skin barrier integrity. Moreover, the authors speculate that *Lactobacillus*, *Staphylococcus*, and *Propionibacterium* (*Cutibacterium*) might act synergistically in skin immunity functions to protect and repair skin from photodamage and inflammation[60].

In our study, it is possible that the skin microbiome is responding to inflammation and AN-associated skin changes in patients with AN as nutritional status improves. Although not significant, we find *Lactobacillus* to increase in patients with AN with weight gain at the inner elbow and lower arm. Conversely, *Lactobacillus* decreases at the forehead, but here, we also find increasing total bacterial load, which significantly correlates with increasing psoriasin concentrations, and we further observe a non-significant increase in *Staphylococcus* at the forehead (Supplementary Fig.3, Supplementary Table S4). The synergistic actions of these three genera to mitigate inflammation and repair skin integrity could perhaps explain why we did not observe significant differences between patients with AN before and after weight gain when analyzing these taxa individually. Indeed, when we visualized the sums of the relative abundances for *Lactobacillus*, *Propionibacterium*, and *Staphylococcus*, we observe a significant difference in relative abundance between patients with AN before and after weight gain and between HC and patients before weight gain. Notably, the sum of these relative abundances in patients with AN after weight gain is not significantly different from HC (Supplementary Fig S4.)

Lastly, our findings contribute to the growing body of evidence demonstrating that BMI significantly associates with skin microbial diversity. In a study of underweight (BMI 15-18.5), normal-weight (BMI 18.5 – 25.0), overweight (25-30), and obese individuals (BMI 30.0-45.0), Brandwein *et al.*[65] found that skin microbial diversity was significantly associated with BMI. Specifically, the authors reported a significant difference in skin microbial diversity between underweight and overweight/ obese individuals and between underweight and normal-weight individuals, but not between normal-weight and overweight/ obese individuals. While limited in terms of sample size and weight categories (underweight and normal-weight only), our work supports the finding that features of the human skin microbiome covary with BMI.

Our study has some limitations. The range of diverse microenvironments (i.e., dry, moist, sebaceous) encompassing human skin as well as the need to consider bacteria living on the skin's surface and those residing within its deeper layers introduce challenges[66–71]. Our sampling strategy included a skin washing method that, to our knowledge, has not been implemented in other skin microbiome surveys. This method is advantageous in that it allows for simultaneous collection of AMPs. A potential downside to this method is that the washing solution likely collects only superficial microbes that can be readily flushed off the skin using the rinsing solution. It is therefore possible that our findings are not necessarily comparable to other surveys

of the skin microbiome in which methods such as skin scraping or swabbing, that can perhaps collect greater numbers of bacteria and bacteria at various depths, were utilized. For example, we find *Propionibacterium* at the oily forehead location in relative abundances less than that typically found in other skin microbiome surveys[72]. This finding may be a result of the rinsing solution not efficiently washing off *Propionibacterium*, which are known to adhere to free fatty acids on the skin[72]. Additionally, it is possible that other bacteria, such as *Lactobacillus*, are more readily washed off the skin's surface and thus may be overrepresented in our results.

Our unique study population is also likely to influence the skin microbial profiles reported here. Our study included young women aged 13 to 18 years with severe and life-threatening anorexia nervosa. Previous surveys found that pre-pubescent children often harbor low levels or no *Propionibacterium* on the skin[73,74]. A common side effect of malnutrition in severe and life-threatening anorexia nervosa is pubertal delay[75]. Thus, it is possible that the relative abundances of *Propionibacterium* in our study reflect those reported in surveys of pre-pubescent children. Moreover, a survey comparing skin microbiota profiles of hands between men and women found that taxa from *Lactobacillaceae* are more abundant on the hands of women[76]. The abundance of *Lactobacillus* reported in our study might reflect a larger phenomenon in which young women harbor greater abundances of these commensal bacteria on the skin compared to men. Future studies, especially those comparing the skin washing method with other established methods in the field, are necessary to verify these hypotheses.

Further, the timing of the second sampling point for AMP and skin microbiota collection may have occurred too soon to sufficiently capture additional meaningful changes in the composition of the skin microbiota. The skin microbiome is remarkably stable at the strain level, despite an ever-changing environment, and the composition of the skin microbiome is largely shaped by host physiology[77]. Given that the patients with AN were still significantly under-weight at the second sampling timepoint, with a mean BMI of 14.54 kg/m^2, it is possible that any immune dysregulation affecting microbial composition at the first timepoint was still present after weight gain. Interestingly, Gibson *et al.*[2] speculate that the pro-inflammatory state in AN is perhaps a primary immunity defect that contributes to the pathogenesis of AN. If immune dysregulation in AN is not necessarily secondary to malnutrition, then it stands to reason improvements in weight and nutrition status in patients with AN would not necessarily affect skin immune processes, and therefore may not lead to substantial changes in skin microbiota. Moreover, it is also feasible that the modest weight gain (at least 2 kg) in patients with AN between sampling timepoints one and two was not enough to alter skin physiology in other ways (e.g., increase sebum production), and therefore not enough to significantly alter microbial community structure. Finally, it is possible that the effect of starvation on AMP levels observed in *Drosophila* is not readily translated to humans. Nutritional status and dietary intake affect human physiological and biochemical processes, yet little is known about the effect of nutrition on human skin physiology[78,79].

In conclusion, this study finds significant statistical correlations between the highly abundant skin AMP psoriasin and features of the skin microbiome of healthy-weight controls compared to patients with AN before and after weight gain. We find no significant statistical correlations between the AMP RNase 7 and skin microbiota. Finally, there is a significant statistical correlation between an individual's BMI and *Lactobacillus*, a significant microbial indicator of health, at all sampling locations. Further studies examining the relationship between caloric and nutritional intake and skin microbiota in the context of eating disorders may help clarify the physiological mechanisms that link nutritional intake with skin physiology.

Methods

Study subjects

The study was approved by the ethics committee of Hannover Medical School (3209-2016) and was conducted following the Declaration of Helsinki and in accordance with relevant guidelines and regulations. All participants or legal guardians provided written informed consent prior to study inclusion.

Thirty-three female patients diagnosed with AN according to DSM-5 criteria[80], and with a body mass index (BMI) of 15 kg/m^2 or below were recruited from two inpatient eating disorder facilities in Germany (Klinik Lüneburger Heide and Hannover Medical School). The DSM-5 defines AN by (a) a restriction of energy intake leading to a significant low body weight, (b) an intense fear of gaining weight or becoming fat, and (c) an unduly influence of body weight or shape on self-worth. Patients with AN were investigated shortly after hospitalization, prior to undergoing an inpatient treatment program to gain weight, and again after having achieved an increase in body cell mass of 2 kg or more. One patient with AN withdrew from the study prior to the second sampling point. Randomly selected healthy-weight control subjects, defined by a BMI between 18.5 and 25.0 kg/m^2, included thirty-three age-matched females without a psychiatric history and free of current mental disorders. Controls were investigated at one time point. Inclusion criteria for all subjects included a minimum age of 16 years, non-smoking status, and to be visually free from skin disorders. All subjects were free from inflammatory disease and immunosuppressive drugs. Final sample sizes included 287 skin rinsing samples from 32 patients with AN and 33 control subjects (Table 1; Supplementary Table S1).

Subjects underwent a clinical interview to gather socio-demographic information and medical history Subjects were weighed using a standardized scale. BMI was calculated using height and weight data. Bioelectrical impedance analysis (BIACORPUS RX 4004 M, Medical Healthcare GmbH, Karlsruhe, Germany) was used to verify an increase in body cell mass of at least 2 kg prior to the second sampling point for the patients with AN.

Sampling procedures

A standardized sampling procedure was implemented by Bendix *et al.*[28], whereby sampling was conducted in the same location by one investigator at the same time to minimize putative influences of the circadian rhythm[81]. Three standardized body sites measuring 1.77 cm^2 in size and comprising diverse skin microenvironments (i.e., sebaceous, moist, and dry body regions) were selected: forehead (sebaceous), inner elbow (moist), and ventral side of the lower forearm (dry; Fig. 1). All subjects avoided cosmetics, lotions, and other topical products the morning of testing. Subjects abstained from physical exercise the morning of sampling days, as exercise may increase AMP expression[82]. Subjects were placed in appropriate positions to collect samples from the various locations. For example, forehead sampling could be carried out with the subject placed in a supine position. Negative sampling controls including ambient air, room controls, and/or negative extraction controls were included for each sampling batch. Ambient air controls containing aliquots of the rinsing buffer solution used at the study site were opened and closed quickly, and then processed as samples. Room controls containing aliquots of the rinsing buffer solution used at the study site were left open for the duration of the rinsing procedure before being processed.

AMPs investigated in this study included psoriasin and RNase 7, which represent the two most abundant AMPs on the surface of human skin[83,84]. AMP data analyzed in this study was previously reported by Bendix *et al.*[28]. AMP sampling was conducted using a skin rinsing method, previously described by Bendix *et al.*[28] (2020), Gläser *et al.*[84], and Wittersheim *et al.*[85]. Briefly, standardized skin sites were washed by pipet with 1 ml of a rinsing buffer solution (10 mM sodium phosphate buffer, pH 7.4 containing 0.1% Triton X-100) using a sterile, DNA-free plastic ring. The buffer solution was collected, centrifuged (10 min, 10.000×g), and diluted 1:10 with 10% (w/v) bovine serum albumin. Samples were stored at − 80 °C until further processing.

Quantitative determination of the AMPs was measured by ELISA, as previously described by Gläser *et al.*[86] and Bendix *et al.*[28]. A monoclonal antibody derived from hybridoma mouse cells was used for the psoriasin ELISA[84]. A polyclonal antibody derived from goat was used for the RNase 7 ELISA[83]. ELISA was performed twice for each sample to ensure reliability. A mean value was calculated from the two sampling measurements and subsequently used in downstream analyses.

Bacteria were collected concurrently with antimicrobial peptides during the skin rinsing protocol and harvested by centrifugation (10 min, 10.000×*g*). Bacterial genomic DNA was extracted from the resulting pellet formed during the centrifugation step of the skin rinsing procedure using the Ultra-Deep Microbiome Prep extraction kit according to the supplier's protocol. Samples were extracted in batches corresponding to collection dates and library preparation and subsequent sequencing (see below) was completed in two batches (Supplementary Table S1).

Bacterial load assessment and 16S rRNA gene sequencing

We adapted ddPCR to measure bacterial loads by targeting the V2 hypervariable region of the 16S rRNA gene, as described by Sze and colleagues[30]. The 20uL ddPCR master mix was prepared according to the manufacturer's protocol with a final primer concentration of 120nM and with 10ng of nucleic acid template. PCR was performed on Bio-Rad C1000 Touch Thermal Cycler with following conditions: 95°C for 5 min, 40 cycles at 95°C for 15 sec and 60°C for 1 min, 4°C for 5 min, 90°C for 5 min, and incubation at 10°C. Final products were transferred to QX200™ Droplet Reader and quantified as gene copies (per 20µL) using Bio-Rad QuantaSoft (v.1.7.4.0917).

16S rRNA amplicon library preparation and sequencing were performed as described in Belheouane *et al.*[87] Briefly, hypervariable regions V1-V2 of the bacterial 16S rRNA gene were amplified and sequencing was performed using the dual-index sequencing strategy for amplicon sequencing on the Illumina MiSeq platform[88]. Negative controls were included in library preparation and sequencing batches. After PCR amplification, the final number of negative controls that were included for sequencing included: ambient air (n = 5), room controls (n = 5), and negative extraction controls with and without rinsing buffer (n = 16, respectively). All controls were processed alongside samples. ZymoBIOMICS Microbial Community Standard cells (Zymo Research) were used as extraction and sequencing controls to assess contamination[36].

Data processing and 16S taxonomic classification

Data processing and statistical analyses were performed in R (version 4.0.5). Processing and taxonomic classification of 16S rRNA gene sequence data was performed as previously described[87]. Sequences were processed using DADA2 (version 1.16.0), resulting in ASV abundance tables[89]. Taxonomic assignment of ASVs was completed in DADA2 with the Bayesian classifier using the NR Silva database training set, version 138[90].

Contamination assessment

As the skin harbors relatively low microbial biomass, the risk of contamination during skin sampling and sample processing is substantial and any contamination introduced during these steps can radically affect data interpretation, as contaminants tend to be preferentially amplified and sequenced over true biological signals within the sample [29,36,91]. Here, we present a detailed description of steps taken throughout this study to mitigate and assess the contribution of contamination.

ddPCR load measurements were used to assess contamination in our dataset via the "frequency" method within the R package "Decontam" (version 1.8.0) in conjunction with Phyloseq (version 1.32.0) [36,37,92]. The strict probability threshold of 0.1 was used and the batch feature within

"Decontam" was utilized to analyze samples according to extraction batch to account for any batch effects and differences in contamination between batches. This resulted in the identification and subsequent removal of 154 ASVs labelled as likely contaminants. To verify ASVs classified as contaminants by the "frequency" method within "Decontam", we visualized five randomly selected contaminant ASVs in frequency plots to view the distribution of the ASV with respect to total bacterial loads obtained from ddPCR. We find that the ASVs identified as contaminants follow the expected pattern in which frequency is inversely proportional to input ddPCR load, as contaminating DNA will account for a larger fraction of the ddPCR load in samples with low biomass.

Next, we utilized the "prevalence" method within "Decontam", in which the prevalence (absence/ presence) of sequence features in samples is compared to the prevalence in negative controls to identify contaminants. The threshold parameter was set to the strict probability threshold of 0.5 and the batch function was utilized to analyze samples according to extraction batch and to account for differences in contamination between batches. An additional 70 ASVs were identified as likely contaminants and subsequently removed from the dataset. Finally, following recommendations of Weyrich *et al.*[38], any ASV belonging to families *Halomonadaceae* (n = 0) and *Shewanellaceae* (n = 14) were removed, as these bacteria represent common contaminants in low biomass samples.

To normalize sequencing coverage after the "Decontam" filtering procedure, we calculated rarefaction curves to determine the sampling threshold; random sub-sampling to 1,000 sequences per sample was performed. Twenty-three samples did not meet the 1,000 sequences coverage threshold and were excluded from the analysis. Additionally, 761 ASVs were removed because they were no longer present in any sample after random sub-sampling.

Ecological and statistical analyses

Alpha diversity was measured using Shannon and Chao1 indices with Vegan (version 2.5-6.) As data were not normally distributed, all three groups were compared against each other using paired (AN before versus AN after) and unpaired (AN before versus HC, AN after versus HC) Wilcoxon rank sum tests. Spearman correlations were performed to assess the relationship between alpha diversity and BMI. Overall differences between groups (beta diversity) were assessed using the Bray-Curtis dissimilarity index in Phyloseq; the vegan package was used to conduct a constrained analysis of principal coordinates ('*capscale'*), a hypothesis-driven ordination that limits the separation of the communities based on the variable tested, for which the 'anova.cca' function was applied to assess significance.

Between group relative abundances at the phylum- and genus-level were calculated in Phyloseq and compared using Wilcoxon rank sum tests. Spearman correlations were performed to assess the relationship between the relative abundances of individual taxa and BMI or concentrations of psoriasin or RNase 7 at individual sampling locations (i.e., forehead, elbow, lower arm). Correction for multiple testing was performed according to Benjamini and Hochberg method[93].

Indicator species analysis was applied using indicspecies (version 1.7.9) with the "r.g." function[94] and 99,999 permutations on a microbial core defined by ASVs classified to the genus-level that are present in at least 5% of all samples. Significant indicator ASVs were selected after correction of *p*-values for multiple testing using the Benjamini and Hochberg method[93]. We additionally calculated Spearman correlations between significant indicator taxa at the genus-level and bacterial loads. In cases of contamination, contaminant features are usually inversely proportional to input DNA concentration, as contaminants tend to be preferentially amplified and sequenced over true biological signals within the sample[29,36,91]. We find no significant negative correlations between bacterial loads and indicator genera reported (see Results), suggesting that these

indicator taxa likely represent true biological signal.

Representative 16S rRNA gene sequences were queried via Ribosomal Database Project SeqMatch (version 3)[95,96]. Results represent classification based on the RDP match score (S_ab), which is the number of unique 7-base oligomers shared between the query sequence and a given RDP sequence for both type- and non-type strains.

Funding sources: This project was funded by the German Centre for Infection Research (DZIF) of the structured doctoral program (StrucMed) at Hannover Medical School (grant to M.B.). Funding was also provided by the German Science Foundation (DFG) grant RTG 1743 "Genes, Environment and Inflammation", the DFG Cluster of Excellence 2167 "Precision Medicine in Chronic Inflammation (PMI)" (grant no. EXC2167 to J.F.B.), and the European Union and its program: ERA-NET (European Research Area Networks) Neuron, Grant/Award Number: 01EW1906A; European Union's Framework Programme for Research and Innovation Horizon 2020; German Ministry for Education and Research (BMBF).

Data availability

The data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB47175 (https://www.ebi.ac.uk/ena/browser/view/PRJEB47175).

Acknowledgements

We would like to thank Katja Cloppenborg-Schmdit, Cornelia Wilgus, Heilwig Hinrichs, and Yasmin Claussen for excellent technical assistance. Thank you to Dr. Malte C. Rühlemann and Shauni Doms for statistical and analytical expertise. Fig. 1a was drawn by Britt M. Hermes and Fig. 1b was drawn by M. Bendix.

Competing interests

The authors state no conflict of interest.

Author contributions

Conceptualization: BMH, JH, MZ, GT, JFB, FR; **Data curation:** BMH (16S rRNA gene data); MB (AMP data); CJ (ddPCR data); **Study (experiment) conduct:** MB, MZ; **Formal analysis:** BMH; **Funding acquisition:** MZ, MB, JH, JFB; **Investigation:** BMH; **Methodology:** BMH, JH, JFB; **Project administration:** BMH, JH, JFB; **Resources:** BMH, JH, JFB; **Software:** BMH; **Supervision:** JH, JFB; **Validation:** BMH, JFB; **Visualization:** BMH; **Writing–Original draft preparation:** BMH; **Writing–Review and editing:** all authors

References

1. Campbell, K. & Peebles, R. Eating Disorders in Children and Adolescents: State of the Art Review. *PEDIATRICS* **134**, 582–592 (2014).
2. Gibson, D. & Mehler, P. S. Anorexia Nervosa and the Immune System—A Narrative Review. *J Clin Med* **8**, (2019).
3. Westmoreland, P., Krantz, M. J. & Mehler, P. S. Medical Complications of Anorexia Nervosa and Bulimia. *Am J Med* **129**, 30–37 (2016).
4. Bourke, C. D., Berkley, J. A. & Prendergast, A. J. Immune Dysfunction as a Cause and Consequence of Malnutrition. *Trends Immunol* **37**, 386–398 (2016).
5. Calder, P. C. & Jackson, A. A. Undernutrition, infection and immune function. *Nutr Res Rev* **13**, 3–29 (2000).
6. Rytter, M. J. H., Kolte, L., Briend, A., Friis, H. & Christensen, V. B. The Immune System in Children with Malnutrition—A Systematic Review. *PLoS One* **9**, e105017 (2014).
7. Bowers, T. K. & Eckert, E. Leukopenia in anorexia nervosa. Lack of increased risk of infection. *Arch Intern Med* **138**, 1520–1523 (1978).
8. Brown, R. F., Bartrop, R. & Birmingham, C. L. Immunological disturbance and infectious disease in anorexia nervosa: a review. *Acta Neuropsychiatr* **20**, 117–128 (2008).
9. Brown, R. F., Bartrop, R., Beumont, P. & Birmingham, C. L. Bacterial infections in anorexia nervosa: delayed recognition increases complications. *Int J Eat Disord* **37**, 261–265 (2005).
10. Nova, E. & Marcos, A. Immunocompetence to assess nutritional status in eating disorders. *Expert Review of Clinical Immunology* **2**, 433–444 (2006).
11. Słotwiński, S. M. & Słotwiński, R. Immune disorders in anorexia. *Cent Eur J Immunol* **42**, 294–300 (2017).
12. Strumìa, R., Varotti, E., Manzato, E. & Gualandi, M. Skin signs in anorexia nervosa. *Dermatology* **203**, 314–317 (2001).
13. Schaible, U. E. & Kaufmann, S. H. E. Malnutrition and Infection: Complex Mechanisms and Global Impacts. *PLoS Med* **4**, (2007).
14. Heilskov, S. *et al.* Dermatosis in children with oedematous malnutrition (Kwashiorkor): a review of the literature. *Journal of the European Academy of Dermatology and Venereology* **28**, 995–1001 (2014).
15. Dalton, B. *et al.* A meta-analysis of cytokine concentrations in eating disorders. *J Psychiatr Res* **103**, 252–264 (2018).
16. Dalton, B. *et al.* Inflammatory Markers in Anorexia Nervosa: An Exploratory Study. *Nutrients* **10**, 1573 (2018).
17. Pasupuleti, M., Schmidtchen, A. & Malmsten, M. Antimicrobial peptides: key components of the innate immune system. *Crit Rev Biotechnol* **32**, 143–171 (2012).
18. Chessa, C. *et al.* Antiviral and Immunomodulatory Properties of Antimicrobial Peptides Produced by Human Keratinocytes. *Front Microbiol* **11**, 1155 (2020).
19. Rademacher, F. *et al.* The Antimicrobial and Immunomodulatory Function of RNase 7 in Skin. *Front Immunol* **10**, 2553 (2019).
20. Lowes, M. A., Suárez-Fariñas, M. & Krueger, J. G. Immunology of Psoriasis. *Annu Rev Immunol* **32**, 227–255 (2014).
21. Watson, P. H., Leygue, E. R. & Murphy, L. C. Psoriasin (S100A7). *The International Journal of Biochemistry & Cell Biology* **30**, 567–571 (1998).
22. Rademacher, F., Simanski, M. & Harder, J. RNase 7 in Cutaneous Defense. *Int J Mol Sci* **17**, 560 (2016).
23. Harder, J. *et al.* Enhanced expression and secretion of antimicrobial peptides in atopic dermatitis and after superficial skin injury. *J Invest Dermatol* **130**, 1355–1364 (2010).
24. Gambichler, T. *et al.* Differential mRNA expression of antimicrobial peptides and proteins in atopic dermatitis as compared to psoriasis vulgaris and healthy skin. *Int Arch Allergy Immunol* **147**, 17–24 (2008).

25. Becker, T. *et al.* FOXO-dependent regulation of innate immune homeostasis. *Nature* **463**, 369–373 (2010).
26. Wu, J., Randle, K. E. & Wu, L. P. ird1 is a Vps15 homologue important for antibacterial immune responses in Drosophila. *Cellular Microbiology* **9**, 1073–1085 (2007).
27. Zinke, I., Schütz, C. S., Katzenberger, J. D., Bauer, M. & Pankratz, M. J. Nutrient control of gene expression in Drosophila: microarray analysis of starvation and sugar-dependent response. *EMBO J* **21**, 6162–6173 (2002).
28. Bendix, M.-C. *et al.* Antimicrobial peptides in patients with anorexia nervosa: comparison with healthy controls and the impact of weight gain. *Sci Rep* **10**, (2020).
29. Eisenhofer, R. *et al.* Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. *Trends in Microbiology* **27**, 105–117 (2019).
30. Sze, M. A., Abbasi, M., Hogg, J. C. & Sin, D. D. A Comparison between Droplet Digital and Quantitative PCR in the Analysis of Bacterial 16S Load in Lung Tissue Samples from Control and COPD GOLD 2. *PLOS ONE* **9**, e110351 (2014).
31. Abellan-Schneyder, I., Schusser, A. J. & Neuhaus, K. ddPCR allows 16S rRNA gene amplicon sequencing of very small DNA amounts from low-biomass samples. *BMC Microbiology* **21**, 349 (2021).
32. Quan, P.-L., Sauzade, M. & Brouzes, E. dPCR: A Technology Review. *Sensors (Basel)* **18**, 1271 (2018).
33. Vogelstein, B. & Kinzler, K. W. Digital PCR. *Proc Natl Acad Sci U S A* **96**, 9236–9241 (1999).
34. Gobert, G. *et al.* Droplet digital PCR improves absolute quantification of viable lactic acid bacteria in faecal samples. *Journal of Microbiological Methods* **148**, 64–73 (2018).
35. Maheshwari, Y., Selvaraj, V., Hajeri, S. & Yokomi, R. Application of droplet digital PCR for quantitative detection of Spiroplasma citri in comparison with real time PCR. *PLOS ONE* **12**, e0184751 (2017).
36. Karstens, L. *et al.* Controlling for Contaminants in Low-Biomass 16S rRNA Gene Sequencing Experiments. *mSystems* **4**, e00290-19, /msystems/4/4/mSys.00290-19.atom (2019).
37. Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A. & Callahan, B. J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* **6**, 226 (2018).
38. Weyrich, L. S. *et al.* Laboratory contamination over time during low-biomass sample analysis. *Molecular Ecology Resources* **19**, 982–996 (2019).
39. Anderson, M. J. & Willis, T. J. Canonical Analysis of Principal Coordinates: A Useful Method of Constrained Ordination for Ecology. *Ecology* **84**, 511–525 (2003).
40. De Caceres, M. & Legendre, P. Associations between species and groups of sites: indices and statistical inference. *Ecology* (2009).
41. Mu, Q., Tavella, V. J. & Luo, X. M. Role of Lactobacillus reuteri in Human Health and Diseases. *Frontiers in Microbiology* **9**, (2018).
42. Delanghe, L. *et al.* The role of lactobacilli in inhibiting skin pathogens. *Biochem Soc Trans* **49**, 617–627 (2021).
43. Kim, J.-H. *et al.* Taxonomic profiling of skin microbiome and correlation with clinical skin parameters in healthy Koreans. *Sci Rep* **11**, 16269 (2021).
44. Kanamoto, T., Terakubo, S. & Nakashima, H. Antimicrobial Susceptibilities of Oral Isolates of Abiotrophia and Granulicatella According to the Consensus Guidelines for Fastidious Bacteria. *Medicines (Basel)* **5**, 129 (2018).
45. Si, J. *et al.* Gut microbiome signatures distinguish type 2 diabetes mellitus from non-alcoholic fatty liver disease. *Computational and Structural Biotechnology Journal* **19**, 5920–5930 (2021).
46. Rath, S., Rud, T., Karch, A., Pieper, D. H. & Vital, M. Pathogenic functions of host microbiota. *Microbiome* **6**, 174 (2018).

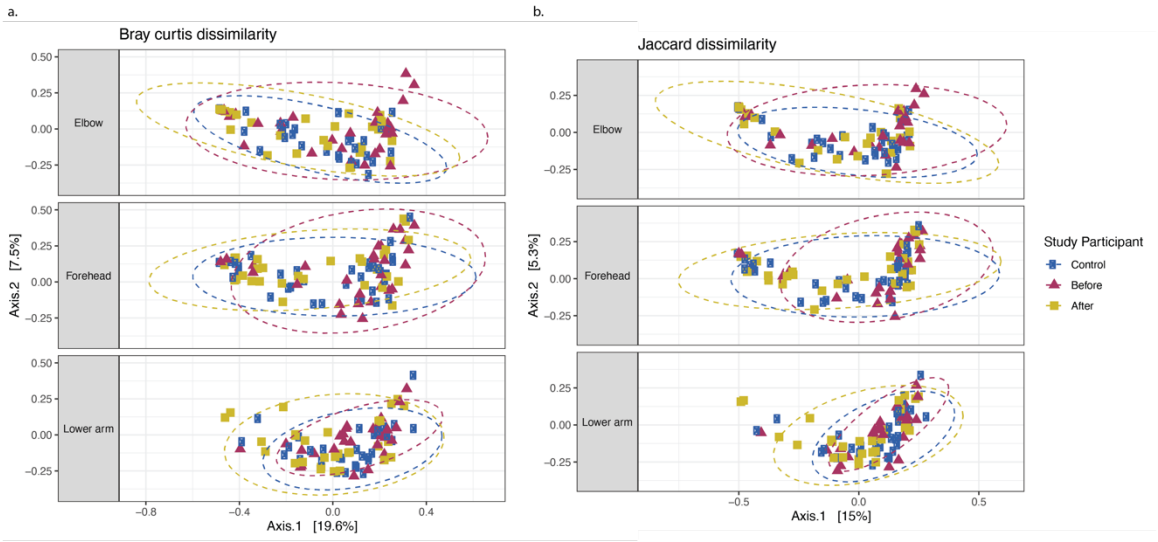
47. Assarsson, M., Söderman, J., Dienus, O. & Seifert, O. Significant Differences in the Bacterial Microbiome of the Pharynx and Skin in Patients with Psoriasis Compared with Healthy Controls. *Acta Derm Venereol* **100**, adv00273 (2020).
48. Omodei, D. *et al.* Immune-metabolic profiling of anorexic patients reveals an anti-oxidant and anti-inflammatory phenotype. *Metabolism* **64**, 396–405 (2015).
49. Chang, H.-W. *et al.* Alteration of the cutaneous microbiome in psoriasis and potential role in Th17 polarization. *Microbiome* **6**, 154 (2018).
50. Alekseyenko, A. V. *et al.* Community differentiation of the cutaneous microbiota in psoriasis. *Microbiome* **1**, 31 (2013).
51. Gläser, R., Köten, B., Wittersheim, M. & Harder, J. Psoriasin: key molecule of the cutaneous barrier? *JDDG: Journal der Deutschen Dermatologischen Gesellschaft* **9**, 897–902 (2011).
52. Nam, B. *et al.* Regulatory effects of Lactobacillus plantarum HY7714 on skin health by improving intestinal condition. *PLoS One* **15**, e0231268 (2020).
53. Jung, Y.-O. *et al.* Lysates of a Probiotic, Lactobacillus rhamnosus, Can Improve Skin Barrier Function in a Reconstructed Human Epidermis Model. *Int J Mol Sci* **20**, (2019).
54. Lee, D. E. *et al.* Clinical Evidence of Effects of Lactobacillus plantarum HY7714 on Skin Aging: A Randomized, Double Blind, Placebo-Controlled Study. *J Microbiol Biotechnol* **25**, 2160–2168 (2015).
55. Markowiak-Kopeć, P. & Śliżewska, K. The Effect of Probiotics on the Production of Short-Chain Fatty Acids by Human Intestinal Microbiome. *Nutrients* **12**, E1107 (2020).
56. Morvan, P.-Y., Vallee, R. & Py, M. Evaluation of the Effects of Stressful Life on Human Skin Microbiota. **4**, 8 (2018).
57. Byeon, J. *et al.* Insights into the skin microbiome of sickle cell disease leg ulcers. *Wound Repair Regen* **29**, 801–809 (2021).
58. Mukherjee, S. *et al.* Sebum and Hydration Levels in Specific Regions of Human Face Significantly Predict the Nature and Diversity of Facial Skin Microbiome. *Sci Rep* **6**, 36062 (2016).
59. Łoś-Rycharska, E. *et al.* A Combined Analysis of Gut and Skin Microbiota in Infants with Food Allergy and Atopic Dermatitis: A Pilot Study. *Nutrients* **13**, 1682 (2021).
60. Li, Z. *et al.* New Insights Into the Skin Microbial Communities and Skin Aging. *Front Microbiol* **11**, 565549 (2020).
61. Kates, A. E., Zimbric, M. L., Mitchell, K., Skarlupka, J. & Safdar, N. The impact of chlorhexidine gluconate on the skin microbiota of children and adults: A pilot study. *Am J Infect Control* **47**, 1014–1016 (2019).
62. Park, H. *et al.* Pilot Study on the Forehead Skin Microbiome and Short Chain Fatty Acids Depending on the SC Functional Index in Korean Cohorts. *Microorganisms* **9**, 2216 (2021).
63. Wang, L. *et al.* Amplicon-based sequencing and co-occurence network analysis reveals notable differences of microbial community structure in healthy and dandruff scalps. *BMC Genomics* **23**, 312 (2022).
64. Kim, H.-J. *et al.* Segregation of age-related skin microbiome characteristics by functionality. *Sci Rep* **9**, 16748 (2019).
65. Brandwein, M., Katz, I., Katz, A. & Kohen, R. Beyond the gut: Skin microbiome compositional changes are associated with BMI. *Human Microbiome Journal* **13**, 100063 (2019).
66. Grice, E. A. *et al.* A diversity profile of the human skin microbiota. *Genome Res* **18**, 1043–1050 (2008).
67. Kim, D. Optimizing methods and dodging pitfalls in microbiome research. 14 (2017).
68. Kong, H. H. Details Matter: Designing Skin Microbiome Studies. *Journal of Investigative Dermatology* **136**, 900–902 (2016).
69. Kong, H. H. *et al.* Performing Skin Microbiome Research: A Method to the Madness. *Journal of Investigative Dermatology* **137**, 561–568 (2017).

70. Meisel, J. S. *et al.* Skin Microbiome Surveys Are Strongly Influenced by Experimental Design. *Journal of Investigative Dermatology* **136**, 947–956 (2016).
71. Nakatsuji, T. *et al.* The microbiome extends to subepidermal compartments of normal skin. *Nat Commun* **4**, 1431 (2013).
72. Grice, E. A. & Segre, J. A. The skin microbiome. *Nat Rev Microbiol* **9**, 244–253 (2011).
73. Mourelatos, K., Eady, E. A., Cunliffe, W. J., Clark, S. M. & Cove, J. H. Temporal changes in sebum excretion and propionibacterial colonization in preadolescent children with and without acne. *Br J Dermatol* **156**, 22–31 (2007).
74. Leyden, J. J., McGinley, K. J., Mills, O. H. & Kligman, A. M. Propionibacterium Levels In Patients With And Without Acne Vulgaris. *Journal of Investigative Dermatology* **65**, 382–384 (1975).
75. Neale, J., Pais, S. M. A., Nicholls, D., Chapman, S. & Hudson, L. D. What Are the Effects of Restrictive Eating Disorders on Growth and Puberty and Are Effects Permanent? A Systematic Review and Meta-Analysis. *J Adolesc Health* **66**, 144–156 (2020).
76. Fierer, N., Hamady, M., Lauber, C. L. & Knight, R. The influence of sex, handedness, and washing on the diversity of hand surface bacteria. *Proceedings of the National Academy of Sciences* **105**, 17994–17999 (2008).
77. Oh, J., Byrd, A. L., Park, M., Kong, H. H. & Segre, J. A. Temporal Stability of the Human Skin Microbiome. *Cell* **165**, 854–866 (2016).
78. Cao, C., Xiao, Z., Wu, Y. & Ge, C. Diet and Skin Aging-From the Perspective of Food Nutrition. *Nutrients* **12**, E870 (2020).
79. Liakou, A. I., Theodorakis, M. J., Melnik, B. C., Pappas, A. & Zouboulis, C. C. Nutritional clinical studies in dermatology. *J Drugs Dermatol* **12**, 1104–1109 (2013).
80. Attia, E. *et al.* Feeding and Eating Disorders in DSM-5. *AJP* **170**, 1237–1239 (2013).
81. Bilska, B. *et al.* Expression of antimicrobial peptide genes oscillates along day/night rhythm protecting mice skin from bacteria. *Experimental Dermatology* **n/a**,.
82. Eda, N., Shimizu, K., Suzuki, S., Lee, E. & Akama, T. Effects of High-Intensity Endurance Exercise on Epidermal Barriers against Microbial Invasion. *J Sports Sci Med* **12**, 44–51 (2013).
83. Köten, B. *et al.* RNase 7 contributes to the cutaneous defense against Enterococcus faecium. *PLoS One* **4**, e6424 (2009).
84. Gläser, R. *et al.* Antimicrobial psoriasin (S100A7) protects human skin from Escherichia coli infection. *Nat Immunol* **6**, 57–64 (2005).
85. Wittersheim, M. *et al.* Differential expression and in vivo secretion of the antimicrobial peptides psoriasin (S100A7), RNase 7, human beta-defensin-2 and -3 in healthy human skin. *Exp Dermatol* **22**, 364–366 (2013).
86. Gläser, R. *et al.* UV-B radiation induces the expression of antimicrobial peptides in human keratinocytes in vitro and in vivo. *J Allergy Clin Immunol* **123**, 1117–1123 (2009).
87. Belheouane, M. *et al.* Assessing similarities and disparities in the skin microbiota between wild and laboratory populations of house mice. *ISME J* (2020) doi:10.1038/s41396-020-0690-7.
88. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform. *Appl. Environ. Microbiol.* **79**, 5112–5120 (2013).
89. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* **13**, 581–583 (2016).
90. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**, D590-596 (2013).
91. Salter, S. J. *et al.* Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biology* **12**, 87 (2014).

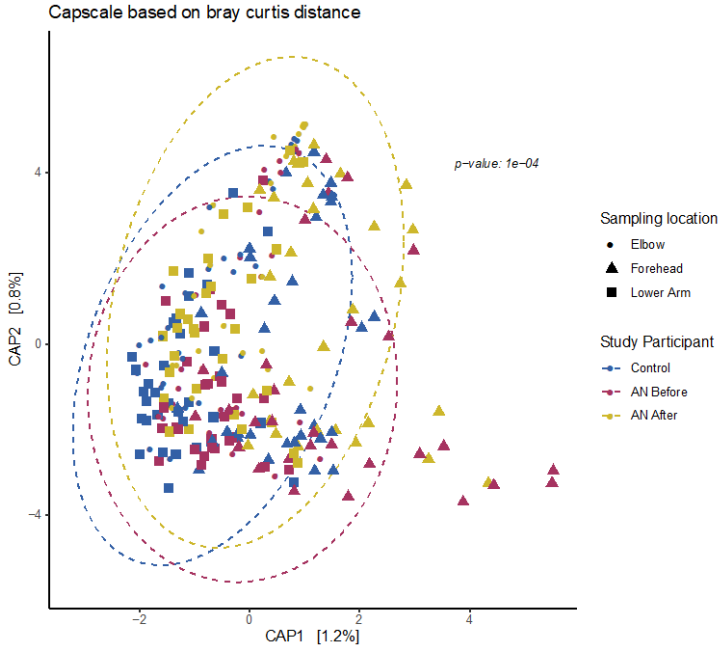
92. McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. in *PloS one* (2013). doi:10.1371/journal.pone.0061217.
93. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 289–300 (1995).
94. Cáceres, M. D. & Legendre, P. Associations between species and groups of sites: indices and statistical inference. *Ecology* **90**, 3566–3574 (2009).
95. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).
96. Cole, J. R. *et al.* The Ribosomal Database Project (RDP-II): sequences and tools for high-throughput rRNA analysis. *Nucleic Acids Res* **33**, D294-296 (2005).

SUPPLEMENTARY FIGURES

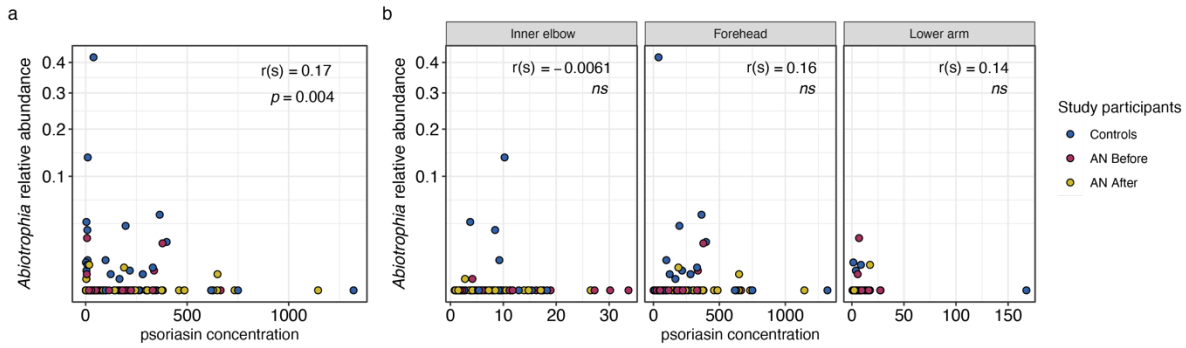
- **Supplementary Figure S1.** Beta diversity
- **Supplementary Figure S2.** Constrained analysis of principal coordinates of the Bray–Curtis dissimilarity index
- **Supplementary Figure S3.** Spearman correlations between indicator *Abiotrophia* and psoriasin concentrations
- **Supplementary Figure S4.** Box plots of relative abundances of *Staphylococcus*, *Propionibacterium*, and the sum relative abundance for *Staphylococcus*, *Propionibacterium*, and *Lactobacillus*



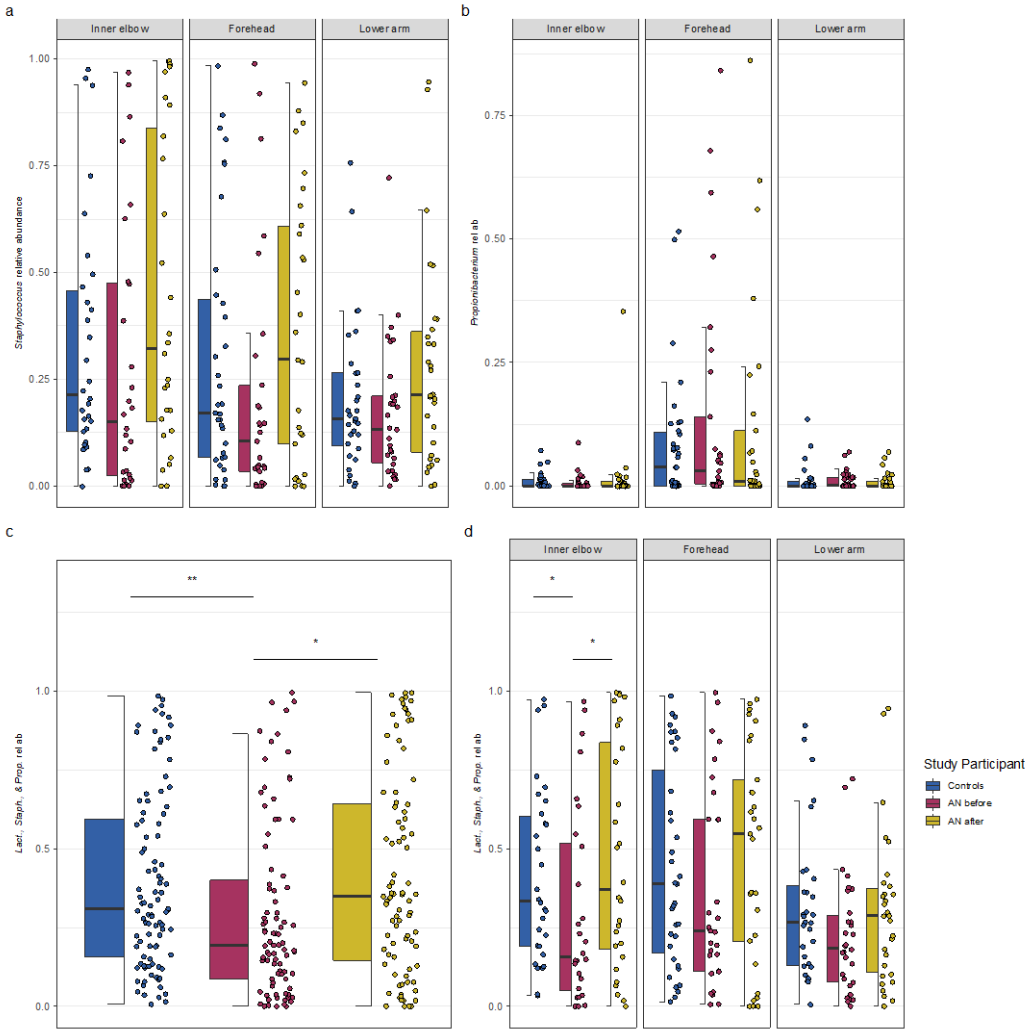
Supplementary Figure 1. Beta diversity based on a. Bray-Curtis dissimilarity index by sampling location and b. Jaccard index by sampling location. Blue represents healthy-weight controls, red represents patients with AN before weight gain, and gold represents patients with AN after weight gain.



Supplementary Figure 2. Constrained analysis of principal coordinates of the Bray–Curtis dissimilarity index with respect to treatment status. Blue represents healthy-weight controls, red represents patients with AN before weight gain, and gold represents patients with AN after weight gain.



Supplementary Fig S3. Spearman correlations between indicator *Abiotrophia* and psoriasin concentrations a. at all locations and b. faceted by sampling location. r(s) = spearman's Rho. AN = anorexia nervosa. ns = not significant. Blue represents healthy-weight controls, red represents patients with AN before weight gain, and gold represents patients with AN after weight gain. *p*-values were adjusted for multiple testing according to Benjamini and Hochberg (1995).



Supplementary Fig S4. Box plots of **a.** *Staphylococcus* and **b.** *Propionibacterium* relative abundances for healthy-weight controls and patients with AN, by weight gain arm (before and after) and, **c.** Sum totals of relative abundances of *Staphylococcus, Propionibacterium, and Lactobacillus* for healthy-weight controls and patients with AN, by weight gain arm (before and after) and, **d.** faceted by sampling location. Wilcoxon test (see Methods**)**; *p*-values: * <0.05; ** < 0.01; *** < 0.001. *p*-values were adjusted for multiple testing according to Benjamini and Hochberg (1995). Line indicates the median concentration; box shows the interquartile range (IQR), and the whiskers are 1.5x IQR. Blue represents healthy-weight controls, red represents patients with AN before weight gain, and gold represents patients with AN after weight gain. Summary statistics provided in Supplementary Table S3. rel = relative; ab = abundance; Lact = *Lactobacillus*; Staph = *Staphylococcus*; Prop = *Propionibacterium*

Supplementary Table S1. Study metadata

Sample_ID	copies	sample_control	control_type	patient_ID	treatment	location	patient_control	pso	R7	age_before	weight_before	weight_change	weight_after	bmi	isolation_date	isolation_batch	sequencing_run
AN-136	22.00	sample	sample	A1	before	S	patient	61.08134493	1.119467	24	39.2			14.2	8/6/2019	batch_19	seq_1
AN-138	70.00	sample	sample	A1	before	U	patient	2.121034786	0.7825305	24	39.2			14.2	8/6/2019	batch_19	seq_1
AN-137	80.00	sample	sample	A1	before	E	patient	30.18449658	0.16675	24	39.2			14.2	8/6/2019	batch_19	seq_2
AN-219	38.00	sample	sample	A1	after	S	patient	198.1783105	1.2443125			5.1	44.3	16.08	10/7/2019	batch_26	seq_2
AN-221	128.00	sample	sample	A1	after	U	patient	8.009266263	0.403375			5.1	44.3	16.08	10/7/2019	batch_26	seq_2
AN-220	148.00	sample	sample	A1	after	E	patient	11.51134538	0.537925			5.1	44.3	16.08	10/7/2019	batch_26	seq_2
AN-227	38.00	sample	sample	A10	before	U	patient	2.23398132	0.70053	26	33.5			10.94	10/8/2019	batch_27	seq_2
AN-226	60.00	sample	sample	A10	before	E	patient	3.817595123	1.12785	26	33.5			10.94	10/8/2019	batch_27	seq_2
AN-225	142.00	sample	sample	A10	before	S	patient	83.98912354	1.09224	26	33.5			10.94	10/8/2019	batch_27	seq_2
AN-252	0.00	sample	sample	A10	after	E	patient	7.17825184	1.375219			6.8	40.3	13.16	11/8/2019	batch_30	seq_1
AN-253	138.00	sample	sample	A10	after	U	patient	4.815585945	0.982145			6.8	40.3	13.16	11/8/2019	batch_30	seq_1
AN-251	290.00	sample	sample	A10	after	S	patient	379.5042506	1.722049			6.8	40.3	13.16	11/8/2019	batch_30	seq_1
AN-5	28.00	sample	sample	A11	after	S	patient	232.2178553	4.2566595			5	42.4	13.84	6/14/2018	batch_2	seq_2
AN-7	36.00	sample	sample	A11	after	U	patient	3.958483709	2.4361695			5	42.4	13.84	6/14/2018	batch_2	seq_2
AN-6	80.00	sample	sample	A11	after	E	patient	4.215227317	3.411432			5	42.4	13.84	6/14/2018	batch_2	seq_2
AN-244	24.00	sample	sample	A11	before	S	patient	27.3562505	1.50239	24	37.4			12.21	10/24/2019	batch_29	seq_1
AN-245	24.00	sample	sample	A11	before	E	patient	3.251965042	0.750925	24	37.4			12.21	10/24/2019	batch_29	seq_2
AN-246	28.00	sample	sample	A11	before	U	patient	1.063736269	0.750925	24	37.4			12.21	10/24/2019	batch_29	seq_2
AN-247	48.00	sample	sample	A12	before	S	patient	162.080837	7.537232	54	37			13.76	10/24/2019	batch_29	seq_1
AN-249	52.00	sample	sample	A12	before	U	patient	6.286633594	1.687366	54	37			13.76	10/24/2019	batch_29	seq_1
AN-248	102.00	sample	sample	A12	before	E	patient	11.47337389	1.583317	54	37			13.76	10/24/2019	batch_29	seq_1
AN-23	26.00	sample	sample	A12	after	U	patient	8.114112085	2.3841555			6.3	43.3	16.1	1/25/2019	batch_5	seq_1
AN-22	50.00	sample	sample	A12	after	E	patient	2.756644777	1.590942			6.3	43.3	16.1	1/25/2019	batch_5	seq_1
AN-21	58.00	sample	sample	A12	after	S	patient	298.4835024	14.3139079			6.3	43.3	16.1	1/25/2019	batch_5	seq_1
AN-261	114.00	sample	sample	A13	before	S	patient	35.81607082	1.560195	25	26			9.55	11/8/2019	batch_30	seq_1
AN-263	62.00	sample	sample	A13	before	U	patient	14.57089209	0.762486	25	26			9.55	11/9/2019	batch_31	seq_1
AN-262	72.00	sample	sample	A13	before	E	patient	14.56407114	1.224926	25	26			9.55	11/9/2019	batch_31	seq_1
AN-27	42.00	sample	sample	A13	after	U	patient	1.923954585	1.2398475			3.2	29.2	10.73	4/16/2019	batch_6	seq_2
AN-26	48.00	sample	sample	A13	after	E	patient	2.4420974	1.018788			3.2	29.2	10.73	4/16/2019	batch_6	seq_1
AN-25	94.00	sample	sample	A13	after	S	patient	649.292711	1.122816			3.2	29.2	10.73	4/16/2019	batch_6	seq_1
AN-2	0.00	sample	sample	A14	before	E	patient	11.72984291	0.8497425	39	28.9			10.24	4/9/2018	batch_1	seq_2
AN-1	30.00	sample	sample	A14	before	S	patient	336.9726529	1.590942	39	28.9			10.24	4/9/2018	batch_1	seq_1
AN-3	60.00	sample	sample	A14	before	U	patient	12.53392456	1.2138405	39	28.9			10.24	4/9/2018	batch_1	seq_2
AN-33	58.00	sample	sample	A14	after	U	patient	17.38640921	15.15969996			6.1	35	12.4	4/16/2019	batch_6	seq_1
AN-32	156.00	sample	sample	A14	after	E	patient	10.54886053	0.6416865			6.1	35	12.4	4/16/2019	batch_6	seq_2
AN-31	336.00	sample	sample	A14	after	S	patient	297.1874548	0.9017565			6.1	35	12.4	4/16/2019	batch_6	seq_1
AN-10	104.00	sample	sample	A15	before	E	patient	1.092521186	0.4596375	19	37.8			14.23	9/20/2018	batch_3	seq_1
AN-11	154.00	sample	sample	A15	before	U	patient	1.058367308	0.4076235	19	37.8			14.23	9/20/2018	batch_3	seq_2
AN-9	156.00	sample	sample	A15	before	S	patient	24.58804528	0.966774	19	37.8			14.23	9/20/2018	batch_3	seq_1
AN-48	17.20	sample	sample	A15	after	S	patient	31.28598507	0.5647375			4.6	42.4	15.96	5/24/2019	batch_9	seq_2
AN-49	17.40	sample	sample	A15	after	E	patient	0.977428156	0.4568625			4.6	42.4	15.96	5/24/2019	batch_9	seq_2
AN-50	24.00	sample	sample	A15	after	U	patient	0.6	0.38135			4.6	42.4	15.96	5/24/2019	batch_9	seq_2
AN-66	16.60	sample	sample	A16	after	E	patient	7.241031929	0.1656			5.9	32.6	12.27	6/6/2019	batch_11	seq_2
AN-67	64.00	sample	sample	A16	after	U	patient	1.689224239	0.7804875			5.9	32.6	12.27	6/6/2019	batch_11	seq_2
AN-65	76.00	sample	sample	A16	after	S	patient	243.9448886	2.2368			5.9	32.6	12.27	6/6/2019	batch_11	seq_1
AN-12	52.00	sample	sample	A16	before	S	patient	200.9265306	1.1618265	45	26.7			10.05	9/20/2018	batch_4	seq_1
AN-13	154.00	sample	sample	A16	before	E	patient	2.888378933	0.706704	45	26.7			10.05	9/20/2018	batch_4	seq_1
AN-14	182.00	sample	sample	A16	before	U	patient	1.654188565	1.69497	45	26.7			10.05	9/20/2018	batch_4	seq_2
AN-16	14.80	sample	sample	A17	before	E	patient	2.3089216	9.55E-02	18	40			14.69	1/25/2019	batch_5	seq_1
AN-17	42.00	sample	sample	A17	before	U	patient	1.67108846	1.3698825	18	40			14.69	1/25/2019	batch_5	seq_1
AN-15	132.00	sample	sample	A17	before	S	patient	485.1602319	3.697509	18	40			14.69	1/25/2019	batch_5	seq_1
AN-38	30.00	sample	sample	A17	after	E	patient	4.151699749	0.706704			5.4	45.4	16.68	4/18/2019	batch_7	seq_1
AN-39	32.00	sample	sample	A17	after	U	patient	8.537455645	1.278858			5.4	45.4	16.68	4/18/2019	batch_7	seq_1
AN-37	46.00	sample	sample	A17	after	S	patient	627.913026	3.411432			5.4	45.4	16.68	4/18/2019	batch_7	seq_1
AN-28	54.00	sample	sample	A18	before	S	patient	237.060761	1.356879	45	37.7			14.19	4/16/2019	batch_6	seq_1
AN-30	110.00	sample	sample	A18	before	U	patient	4.737527698	1.95504	45	37.7			14.19	4/16/2019	batch_6	seq_2
AN-29	112.00	sample	sample	A18	before	E	patient	17.48140271	0.810732	45	37.7			14.19	4/16/2019	batch_6	seq_2
AN-42	26.00	sample	sample	A18	after	U	patient	4.099129126	3.3479125			4.6	42.3	15.92	4/18/2019	batch_7	seq_1
AN-40	50.00	sample	sample	A18	after	S	patient	176.2884151	4.1893375			4.6	42.3	15.92	4/18/2019	batch_7	seq_1
AN-41	120.00	sample	sample	A18	after	E	patient	8.500611401	1.33065			4.6	42.3	15.92	4/18/2019	batch_7	seq_1
AN-107	28.00	sample	sample	A19	after	U	patient	1.198371566	0.409775			4.3	30	10.76	6/21/2019	batch_15	seq_1
AN-105	42.00	sample	sample	A19	after	S	patient	2.903080703	0.409775			4.3	30	10.76	6/21/2019	batch_15	seq_1
AN-106	50.00	sample	sample	A19	after	E	patient	1.980158915	1.46375			4.3	30	10.76	6/21/2019	batch_15	seq_1
AN-35	17.40	sample	sample	A19	before	E	patient	1.552523629	0.264585	21	25.7			9.22	4/18/2019	batch_7	seq_2
AN-36	24.00	sample	sample	A19	before	U	patient	1.599770956	0.524655	21	25.7			9.22	4/18/2019	batch_7	seq_2
AN-34	52.00	sample	sample	A19	before	S	patient	1.638259888	0	21	25.7			9.22	4/18/2019	batch_7	seq_1
AN-141	20.00	sample	sample	A2	before	U	patient	1.164829128	0.9451895	20	40.5			14.35	8/6/2019	batch_19	seq_1
AN-140	32.00	sample	sample	A2	before	E	patient	4.751441179	3.2921265	20	40.5			14.35	8/6/2019	batch_19	seq_1
AN-139	358.00	sample	sample	A2	before	S	patient	664.6484404	23.51517481	20	40.5			14.35	8/6/2019	batch_19	seq_1
AN-216	60.00	sample	sample	A2	after	E	patient	5.17808065	3.87925			3.7	44.2	15.85	10/7/2019	batch_26	seq_2
AN-217	70.00	sample	sample	A2	after	U	patient	3.984147773	2.556175			3.7	44.2	15.85	10/7/2019	batch_26	seq_2
AN-215	3420.00	sample	sample	A2	after	S	patient	1144.74906	10.24322273			3.7	44.2	15.85	10/7/2019	batch_26	seq_2
AN-54	22.00	sample	sample	A20	before	U	patient	8.616575497	4.52375	18	40.7			13.95	6/4/2019	batch_10	seq_1
AN-52	74.00	sample	sample	A20	before	S	patient	333.225851	5.303357508	18	40.7			13.95	6/4/2019	batch_10	seq_2
AN-53	76.00	sample	sample	A20	before	E	patient	18.93526601	1.7945125	18	40.7			13.95	6/4/2019	batch_10	seq_2
AN-102	24.00	sample	sample	A20	after	S	patient	459.8573917	5.36804			6.3	47	16.08	6/21/2019	batch_15	seq_1
AN-104	44.00	sample	sample	A20	after	U	patient	7.04352398	1.31711			6.3	47	16.08	6/21/2019	batch_15	seq_1
AN-103	66.00	sample	sample	A20	after	E	patient	26.51117875	0.52892			6.3	47	16.08	6/21/2019	batch_15	seq_1
AN-56	18.00	sample	sample	A21	before	E	patient	11.17839466	0.3921375	23	34			11.36	6/4/2019	batch_10	seq_2
AN-57	82.00	sample	sample	A21	before	U	patient	1.627764472	0.2626875	23	34			11.36	6/4/2019	batch_10	seq_2
AN-55	742.00	sample	sample	A21	before	S	patient	115.4194617	0.6510375	23	34			11.36	6/4/2019	batch_10	seq_1
AN-85	30.00	sample	sample	A21	after	S	patient	50.00856475	0.16232			8.7	42.7	14.27	6/14/2019	batch_13	seq_1
AN-87	86.00	sample	sample	A21	after	U	patient	1.828043992	0.41894			8.7	42.7	14.27	6/14/2019	batch_13	seq_1
AN-86	128.00	sample	sample	A21	after	E	patient	1.47162503	3.40E-02			8.7	42.7	14.27	6/14/2019	batch_13	seq_1
AN-335	184.00	sample	sample	A22	after	S	patient	303.69188	3.078588			7.1	42.2	14.43	9/28/2020	batch_38	seq_2
AN-337	734.00	sample	sample	A22	after	U	patient	14.576714	0.738569			7.1	42.2	14.43	9/28/2020	batch_38	seq_2
AN-336	1582.00	sample	sample	A22	after	E	patient	14.78284462	1.3670865			7.1	42.2	14.43	9/28/2020	batch_38	seq_2
AN-46	58.00	sample	sample	A22	before	U	patient	5.575329432	0.7697	29	35.1			12	5/22/2019	batch_8	seq_1
AN-44	128.00	sample	sample	A22	before	S	patient	193.7106348	1.4924625	29	35.1			12	5/22/2019	batch_8	seq_1
AN-45	496.00	sample	sample	A22	before	E	patient	33.65765485	1.2335625	29	35.1			12	5/22/2019	batch_8	seq_1
AN-60	14.20	sample	sample	A23	before	U	patient	3.747543416	0.446075	23	35			12.11	6/4/2019	batch_10	seq_2
AN-59	158.00	sample	sample	A23	before	E	patient	11.75632796	0.8452125	23	35			12.11	6/4/2019	batch_10	seq_2
AN-58	858.00	sample	sample	A23	before	S	patient	330.4557164	1.3845875	23	35			12.11	6/4/2019	batch_10	seq_1
AN-94	28.00	sample	sample	A23	after	E	patient	17.07584784	0.739715			8.5	43.5	14.88	6/20/2019	batch_14	seq_1
AN-95	96.00	sample	sample	A23	after	U	patient	2.266091306	1.1888			8.5	43.5	14.88	6/20/2019	batch_14	seq_1
AN-93	130.00	sample	sample	A23	after	S	patient	189.1128888	0.69389			8.5	43.5	14.88	6/20/2019	batch_14	seq_1
AN-69	26.00	sample	sample	A24	before	E	patient	34.80977917	0.4784375	24	32.7			11.05	6/6/2019	batch_11	seq_2
AN-70	36.00	sample	sample	A24	before	U	patient	27.22779743	0.856	24	32.7			11.05	6/6/2019	batch_11	seq_1
AN-68	54.00	sample	sample	A24	before	S	patient	41.69101856	0.3489875	24	32.7			11.05	6/6/2019	batch_11	seq_1
AN-290	38.00	sample	sample	A24	after	U	patient	0.844249313	0.524089			5.3	38	12.84	9/18/2020	batch_34	seq_2
AN-289	66.00	sample	sample	A24	after	E	patient	6.473797463	0.852808			5.3	38	12.84	9/18/2020	batch_34	seq_2
AN-288	74.00	sample	sample	A24	after	S	patient	21.20767044	1.095022			5.3	38	12.84	9/18/2020	batch_34	seq_2
AN-78	50.00	sample	sample	A25	before	U	patient	9.387683492	0.748125	19	37.5			12.82	6/7/2019	batch_12	seq_2
AN-77	104.00	sample	sample	A25	before	E	patient	2.238915883	9.01E-02	19	37.5			12.82	6/7/2019	batch_12	seq_2
AN-76	118.00	sample	sample	A25	before	S	patient	181.4298994	0.7589125	19	37.5			12.82	6/7/2019	batch_12	seq_1
AN-297	46.00	sample	sample	A25	after	U	patient	5.251536626	2.063878			5.6	43.1	14.74	9/18/2020	batch_34	seq_2
AN-296	112.00	sample	sample	A25	after	E	patient	1.111596705	0.19537			5.6	43.1	14.74	9/18/2020	batch_34	seq_2
AN-295	144.00	sample	sample	A25	after	S	patient	306.6966126	1.383372			5.6	43.1	14.74	9/18/2020	batch_34	seq_2
AN-83	70.00	sample	sample	A26	before	E	patient	1.826941488	0.501425	18	31.8			13.24	6/14/2019	batch_13	seq_1

AN-84	116.00	sample	sample	A26	before	U	patient	3.378643196	1.3721	18	31.8			13.24	6/14/2019	batch_13	seq_1
AN-82	672.00	sample	sample	A26	before	S	patient	223.2310416	0.76721	18	31.8			13.24	6/14/2019	batch_13	seq_1
AN-330	42.00	sample	sample	A26	after	U	patient	16.62271615	0.989976			7.1	38.9	16.19	9/25/2020	batch_37	seq_2
AN-329	130.00	sample	sample	A26	after	E	patient	1.955426806	0.5935265			7.1	38.9	16.19	9/25/2020	batch_37	seq_2
AN-328	862.00	sample	sample	A26	after	S	patient	142.8909272	1.2123745			7.1	38.9	16.19	9/25/2020	batch_37	seq_2
AN-96	30.00	sample	sample	A27	before	S	patient	50.25345046	0.30896	52	37.6			14.15	6/20/2019	batch_14	seq_1
AN-98	66.00	sample	sample	A27	before	U	patient	16.25873456	1.39043	52	37.6			14.15	6/20/2019	batch_14	seq_1
AN-97	146.00	sample	sample	A27	before	E	patient	27.24826311	1.7387	52	37.6			14.15	6/20/2019	batch_14	seq_1
AN-298	374.00	sample	sample	A27	after	S	patient	23.42363048	1.037352			5.1	42.7	16.07	9/18/2020	batch_34	seq_2
AN-300	784.00	sample	sample	A27	after	U	patient	5.275754333	0.489487			5.1	42.7	16.07	9/18/2020	batch_34	seq_2
AN-299	862.00	sample	sample	A27	after	E	patient	4.035425635	0.59906			5.1	42.7	16.07	9/18/2020	batch_34	seq_2
AN-117	32.00	sample	sample	A28	before	U	patient	1.693145211	0.824436	26	35.9			12.87	6/25/2019	batch_16	seq_1
AN-116	46.00	sample	sample	A28	before	E	patient	9.202690461	0.741124	26	35.9			12.87	6/25/2019	batch_16	seq_1
AN-115	890.00	sample	sample	A28	before	S	patient	351.0057974	0.793194	26	35.9			12.87	6/25/2019	batch_16	seq_1
AN-333	330.00	sample	sample	A28	after	U	patient	15.666565	0.351789			5	40.9	14.67	9/25/2020	batch_37	seq_2
AN-332	464.00	sample	sample	A28	after	E	patient	3.638695947	0.641874			5	40.9	14.67	9/25/2020	batch_37	seq_2
AN-331	540.00	sample	sample	A28	after	S	patient	647.4242298	0.9416285			5	40.9	14.67	9/25/2020	batch_37	seq_2
AN-282	0.00	sample	sample	A29	before	E	patient	0.799461287	0.316477	26	37.3			13.22	9/15/2020	batch_33	seq_2
AN-281	74.00	sample	sample	A29	before	S	patient	21.6538234	0.333778	26	37.3			13.22	9/15/2020	batch_33	seq_2
AN-283	202.00	sample	sample	A29	before	U	patient	0.904660825	0.287642	26	37.3			13.22	9/15/2020	batch_33	seq_2
AN-321	314.00	sample	sample	A29	after	S	patient	30.52420639	0.2865435			5.2	42.5	15.06	9/24/2020	batch_36	seq_2
AN-323	710.00	sample	sample	A29	after	U	patient	5.575900903	4.78E-02			5.2	42.5	15.06	9/24/2020	batch_36	seq_2
AN-322	1016.00	sample	sample	A29	after	E	patient	1.436746491	6.23E-02			5.2	42.5	15.06	9/24/2020	batch_36	seq_2
AN-154	32.00	sample	sample	A3	before	S	patient	23.83928845	0.6612625	21	31			11.25	8/15/2019	batch_20	seq_1
AN-156	38.00	sample	sample	A3	before	U	patient	2.712538233	2.71315	21	31			11.25	8/16/2019	batch_21	seq_1
AN-155	82.00	sample	sample	A3	before	E	patient	0.979003883	0.582775	21	31			11.25	8/16/2019	batch_21	seq_1
AN-168	26.00	sample	sample	A3	after	E	patient	9.06180725	2.30298			6.3	37.3	13.54	9/3/2019	batch_22	seq_1
AN-167	28.00	sample	sample	A3	after	S	patient	31.95063464	1.80444			6.3	37.3	13.54	9/3/2019	batch_22	seq_1
AN-169	36.00	sample	sample	A3	after	U	patient	3.227698163	1.10411			6.3	37.3	13.54	9/3/2019	batch_22	seq_2
AN-287	26.00	sample	sample	A30	before	U	patient	6.88647825	2.300325	18	38.9			14.46	9/18/2020	batch_34	seq_2
AN-286	124.00	sample	sample	A30	before	E	patient	16.99077262	0.795138	18	38.9			14.46	9/18/2020	batch_34	seq_2
AN-285	126.00	sample	sample	A30	before	S	patient	211.6983986	1.268032	18	38.9			14.46	9/18/2020	batch_34	seq_2
AN-350	412.00	sample	sample	A30	after	U	patient	9.877673526	0.9803065			8.3	47.2	17.55	10/2/2020	batch_39	seq_2
AN-349	508.00	sample	sample	A30	after	E	patient	11.56863805	1.2703915			8.3	47.2	17.55	10/2/2020	batch_39	seq_2
AN-348	560.00	sample	sample	A30	after	S	patient	186.1050228	0.2841025			8.3	47.2	17.55	10/2/2020	batch_39	seq_2
AN-306	36.00	sample	sample	A31	before	E	patient	17.10367716	4.779168	19	37.9			13.59	9/22/2020	batch_35	seq_2
AN-307	96.00	sample	sample	A31	before	U	patient	4.764426759	4.5838365	19	37.9			13.59	9/22/2020	batch_35	seq_2
AN-305	116.00	sample	sample	A31	before	S	patient	163.0776168	0.1563225	19	37.9			13.59	9/22/2020	batch_35	seq_2
AN-340	206.00	sample	sample	A31	after	U	patient	17.371518	2.401723			4.8	42.7	15.31	9/28/2020	batch_38	seq_2
AN-339	452.00	sample	sample	A31	after	E	patient	4.394140094	1.164027			4.8	42.7	15.31	9/28/2020	batch_38	seq_2
AN-338	864.00	sample	sample	A31	after	S	patient	486.5022798	0.854603			4.8	42.7	15.31	9/28/2020	batch_38	seq_2
AN-308	266.00	sample	sample	A32	before	S	patient	378.4234059	1.0534005	19	34.7			11.59	9/22/2020	batch_35	seq_2
AN-309	1176.00	sample	sample	A32	before	E	patient	5.991479831	2.9343705	19	34.7			11.59	9/22/2020	batch_35	seq_2
AN-310	1840.00	sample	sample	A32	before	U	patient	1.950807485	2.1241065	19	34.7			11.59	9/22/2020	batch_35	seq_2
AN-353	278.00	sample	sample	A32	after	U	patient	1.28783318	2.35158			5.4	40.1	13.4	10/2/2020	batch_39	seq_2
AN-351	674.00	sample	sample	A32	after	S	patient	374.0891655	1.172805			5.4	40.1	13.4	10/2/2020	batch_39	seq_2
AN-352	752.00	sample	sample	A32	after	E	patient	7.822609655	1.857255			5.4	40.1	13.4	10/2/2020	batch_39	seq_2
AN-320	506.00	sample	sample	A33	before	U	patient	1.86142976	0.40953	17	44.3			14.8	9/24/2020	batch_36	seq_2
AN-319	642.00	sample	sample	A33	before	E	patient	2.040185887	1.393422	17	44.3			14.8	9/24/2020	batch_36	seq_2
AN-318	682.00	sample	sample	A33	before	S	patient	126.9571349	7.67E-02	17	44.3			14.8	9/24/2020	batch_36	seq_2
AN-157	24.00	sample	sample	A4	before	S	patient	17.0661737	0.268825	27	34.4			12.64	8/16/2019	batch_21	seq_2
AN-158	34.00	sample	sample	A4	before	E	patient	3.758971225	0.7846	27	34.4			12.64	8/16/2019	batch_21	seq_1
AN-159	72.00	sample	sample	A4	before	U	patient	2.011276225	0.537925	27	34.4			12.64	8/16/2019	batch_21	seq_1
AN-171	18.00	sample	sample	A4	after	E	patient	3.508847657	0.58183			3.1	37.5	13.77	9/3/2019	batch_22	seq_1
AN-170	28.00	sample	sample	A4	after	S	patient	41.28027145	0.85484			3.1	37.5	13.77	9/3/2019	batch_22	seq_1
AN-172	80.00	sample	sample	A4	after	U	patient	1.941931017	0.67679			3.1	37.5	13.77	9/3/2019	batch_22	seq_1
AN-181	22.00	sample	sample	A5	before	E	patient	1.908545513	0.9079375	20	41.6			14.23	9/12/2019	batch_23	seq_2
AN-182	28.00	sample	sample	A5	before	U	patient	6.842517603	1.8946375	20	41.6			14.23	9/12/2019	batch_23	seq_1
AN-180	116.00	sample	sample	A5	before	S	patient	35.71168365	0.3697375	20	41.6			14.23	9/12/2019	batch_23	seq_1
AN-230	42.00	sample	sample	A5	after	U	patient	7.645265921	2.35046			3.9	45.5	15.56	10/8/2019	batch_27	seq_2
AN-228	74.00	sample	sample	A5	after	S	patient	49.7910452	0.60557			3.9	45.5	15.56	10/8/2019	batch_27	seq_2
AN-229	1002.00	sample	sample	A5	after	E	patient	8.117044453	3.16949			3.9	45.5	15.56	10/8/2019	batch_27	seq_2
AN-200	108.00	sample	sample	A6	before	E	patient	4.159491173	0.358525	19	23.4			9.26	9/30/2019	batch_24	seq_1
AN-199	244.00	sample	sample	A6	before	S	patient	280.9696544	0.4594375	19	23.4			9.26	9/30/2019	batch_24	seq_1
AN-201	570.00	sample	sample	A6	before	U	patient	1.238753505	0.3697375	19	23.4			9.26	9/30/2019	batch_24	seq_1
AN-18	32.00	sample	sample	A6	after	S	patient	733.548114	0.992781			5.7	29.1	11.51	1/25/2019	batch_5	seq_2
AN-20	38.00	sample	sample	A6	after	U	patient	2.392143829	0.862746			5.7	29.1	11.51	1/25/2019	batch_5	seq_2
AN-19	76.00	sample	sample	A6	after	E	patient	8.120503154	0.3816165			5.7	29.1	11.51	1/25/2019	batch_5	seq_2
AN-151	34.00	sample	sample	A7	before	S	patient	72.49733377	3.84608	33	30.1			10.79	8/15/2019	batch_20	seq_2
AN-153	34.00	sample	sample	A7	before	U	patient	25.80575849	1.02102	33	30.1			10.79	8/15/2019	batch_20	seq_1
AN-152	58.00	sample	sample	A7	before	E	patient	4.759811031	0.73614	33	30.1			10.79	8/15/2019	batch_20	seq_1
AN-235	20.00	sample	sample	A7	after	S	patient	87.2397401	2.16054			9.1	30.1	14.06	10/13/2019	batch_28	seq_2
AN-236	24.00	sample	sample	A7	after	E	patient	4.436218645	3.75112			9.1	30.1	14.06	10/13/2019	batch_28	seq_2
AN-237	40.00	sample	sample	A7	after	U	patient	11.38979818	6.25569			9.1	30.1	14.06	10/13/2019	batch_28	seq_1
AN-188	30.00	sample	sample	A8	before	U	patient	12.45175561	3.097808	21	45.5			13.15	9/23/2019	batch_23	seq_2
AN-186	34.00	sample	sample	A8	before	S	patient	146.8366871	1.641122	21	45.5			13.15	9/23/2019	batch_23	seq_1
AN-187	34.00	sample	sample	A8	before	E	patient	16.22111532	4.357957	21	45.5			13.15	9/23/2019	batch_23	seq_2
AN-266	22.00	sample	sample	A8	after	U	patient	4.57910845	1.24655			6.8	45.5	15.12	11/9/2019	batch_31	seq_1
AN-264	206.00	sample	sample	A8	after	S	patient	98.89217901	0.90232			6.8	45.5	15.12	11/9/2019	batch_31	seq_1
AN-265	714.00	sample	sample	A8	after	E	patient	3.774355243	1.34151			6.8	45.5	15.12	11/9/2019	batch_31	seq_1
AN-222	36.00	sample	sample	A9	before	S	patient	58.03191373	1.7807	20	36.9			14.41	10/8/2019	batch_27	seq_2
AN-223	42.00	sample	sample	A9	before	E	patient	2.718029274	1.02102	20	36.9			14.41	10/8/2019	batch_27	seq_2
AN-224	48.00	sample	sample	A9	before	U	patient	1.852563863	1.63826	20	36.9			14.41	10/8/2019	batch_27	seq_2
AN-273	76.00	sample	sample	A9	after	U	patient	1.952487425	0.87858			5.2	42.1	16.45	11/10/2019	batch_32	seq_1
AN-272	96.00	sample	sample	A9	after	E	patient	13.88642349	0.79549			5.2	42.1	16.45	11/10/2019	batch_32	seq_1
AN-271	290.00	sample	sample	A9	after	S	patient	219.5813227	1.0685			5.2	42.1	16.45	11/10/2019	batch_32	seq_1
AN-146	17.40	control	blank_control	blank											8/15/2019	batch_20	seq_1
AN-354	0.00	control	blank_control	blank											10/2/2020	batch_39	seq_2
AN-355	0.00	control	blank_control	blank											10/2/2020	batch_39	seq_2
AN-367	22.00	control	blank_control	blank											10/7/2020	batch_40	seq_2
AN-108	94.00	control	buffer_control	buffer_control											6/21/2019	batch_15	seq_1
AN-145	42.00	control	buffer_control	buffer_control											8/6/2019	batch_19	seq_1
AN-231	28.00	control	buffer_control	buffer_control											10/8/2019	batch_27	seq_2
AN-280	28.00	control	buffer_control	buffer_control											11/10/2019	batch_32	seq_1
AN-142	24.00	sample	sample	CA1	mcontrol	S	control	96.8483763	0.4107385	25	67.5			22.55	8/6/2019	batch_19	seq_1
AN-143	34.00	sample	sample	CA1	mcontrol	E	control	1.866044731	0.9219525	25	67.5			22.55	8/6/2019	batch_19	seq_1
AN-144	50.00	sample	sample	CA1	mcontrol	U	control	1.569994171	0.561779	25	67.5			22.55	8/6/2019	batch_19	seq_1
AN-184	34.00	sample	sample	CA10	mcontrol	E	control	9.693765695	0.92434	21	60			22.58	9/23/2019	batch_23	seq_1
AN-185	36.00	sample	sample	CA10	mcontrol	U	control	1.80775154	0.635315	21	60			22.58	9/23/2019	batch_23	seq_1
AN-183	184.00	sample	sample	CA10	mcontrol	S	control	217.9424313	0.935901	21	60			22.58	9/23/2019	batch_23	seq_1
AN-190	24.00	sample	sample	CA11	mcontrol	E	control	6.372307229	0.80873	22	57			21.72	9/23/2019	batch_23	seq_2
AN-189	26.00	sample	sample	CA11	mcontrol	S	control	81.4152643	2.450392	22	57			21.72	9/23/2019	batch_23	seq_1
AN-191	46.00	sample	sample	CA11	mcontrol	U	control	0.962460428	0.496583	22	57			21.72	9/23/2019	batch_23	seq_2
AN-234	28.00	sample	sample	CA12	mcontrol	U	control	1.514613509	3.53746	23	62			21.71	10/13/2019	batch_28	seq_2
AN-233	40.00	sample	sample	CA12	mcontrol	E	control	1.845190227	1.74509	23	62			21.71	10/13/2019	batch_28	seq_2
AN-232	58.00	sample	sample	CA12	mcontrol	S	control	31.84711429	1.75696	23	62			21.71	10/13/2019	batch_28	seq_2
AN-240	34.00	sample	sample	CA13	mcontrol	U	control	1.537192187	1.41273	21	64			20.2	10/13/2019	batch_28	seq_2
AN-238	142.00	sample	sample	CA13	mcontrol	S	control	35.29086889	2.14867	21	64			20.2	10/13/2019	batch_28	seq_1
AN-241	50.00	sample	sample	CA14	mcontrol	S	control	166.6886304	2.02997	22	66.7			23.63	10/13/2019	batch_28	seq_1
AN-242	34.00	sample	sample	CA14	mcontrol	E	control	7.33713773	1.56704	22	66.7			23.63	10/24/2019	batch_29	seq_1
AN-243	48.00	sample	sample	CA14	mcontrol	U	control	4.396142495	4.85503	22	66.7			23.63	10/24/2019	batch_29	seq_2

AN-268	46.00	sample	sample	CA15	mcontrol	E	control	17.54012906	1.51956	26	64.5			25.84	11/9/2019	batch_31	seq_1
AN-269	62.00	sample	sample	CA15	mcontrol	U	control	5.305518695	1.49582	26	64.5			25.84	11/9/2019	batch_31	seq_1
AN-267	378.00	sample	sample	CA15	mcontrol	S	control	345.2072019	5.35357	26	64.5			25.84	11/9/2019	batch_31	seq_1
AN-275	28.00	sample	sample	CA16	mcontrol	E	control	2.715662342	1.421463	26	56.9			20.9	11/10/2019	batch_32	seq_1
AN-276	166.00	sample	sample	CA16	mcontrol	U	control	1.376266998	0.681559	26	56.9			20.9	11/10/2019	batch_32	seq_1
AN-274	402.00	sample	sample	CA16	mcontrol	S	control	302.8394061	3.733663	26	56.9			20.9	11/10/2019	batch_32	seq_1
AN-278	46.00	sample	sample	CA17	mcontrol	E	control	5.726954336	0.993706	23	62			21.71	11/10/2019	batch_32	seq_1
AN-279	122.00	sample	sample	CA17	mcontrol	U	control	3.490932477	1.305853	23	62			21.71	11/10/2019	batch_32	seq_1
AN-277	648.00	sample	sample	CA17	mcontrol	S	control	196.777015	0.80873	23	62			21.71	11/10/2019	batch_32	seq_1
AN-256	62.00	sample	sample	CA18	mcontrol	U	control	8.669736918	2.057318	22	62.5			22.14	11/8/2019	batch_30	seq_1
AN-255	70.00	sample	sample	CA18	mcontrol	E	control	9.986780575	2.092001	22	62.5			22.14	11/8/2019	batch_30	seq_1
AN-254	324.00	sample	sample	CA18	mcontrol	S	control	123.3068174	1.537073	22	62.5			22.14	11/8/2019	batch_30	seq_1
AN-258	56.00	sample	sample	CA19	mcontrol	E	control	8.451928504	1.282731	30	61			21.61	11/8/2019	batch_30	seq_1
AN-257	60.00	sample	sample	CA19	mcontrol	S	control	398.8657609	1.976391	30	61			21.61	11/8/2019	batch_30	seq_1
AN-259	86.00	sample	sample	CA19	mcontrol	U	control	8.765042579	1.028389	30	61			21.61	11/8/2019	batch_30	seq_1
AN-160	36.00	sample	sample	CA2	mcontrol	S	control	57.33052925	1.0537	22	69			21.53	8/16/2019	batch_21	seq_1
AN-162	40.00	sample	sample	CA2	mcontrol	U	control	1.205410433	1.3564375	22	69			21.53	8/16/2019	batch_21	seq_1
AN-161	72.00	sample	sample	CA2	mcontrol	E	control	9.441892605	1.1434	22	69			21.53	8/16/2019	batch_21	seq_1
AN-211	66.00	sample	sample	CA20	mcontrol	U	control	7.835033767	3.040003	57	67			24.02	10/1/2019	batch_25	seq_1
AN-209	220.00	sample	sample	CA20	mcontrol	S	control	330.7661648	1.143999	57	67			24.02	10/1/2019	batch_25	seq_1
AN-210	350.00	sample	sample	CA20	mcontrol	E	control	6.718302565	3.69898	57	67			24.02	10/1/2019	batch_25	seq_1
AN-79	364.00	sample	sample	CA21	mcontrol	S	control	217.403798	0.65469	36	60			21.01	6/7/2019	batch_12	seq_1
AN-81	36.00	sample	sample	CA21	mcontrol	U	control	1.755549997	0.6937005	36	60			21.01	6/14/2019	batch_13	seq_1
AN-80	52.00	sample	sample	CA21	mcontrol	E	control	5.39183264	1.0317915	36	60			21.01	6/14/2019	batch_13	seq_2
AN-61	42.00	sample	sample	CA22	mcontrol	S	control	750.1712626	3.8765	46	57			21.45	6/6/2019	batch_11	seq_1
AN-64	58.00	sample	sample	CA22	mcontrol	U	control	0.727422422	0.5971	46	57			21.45	6/6/2019	batch_11	seq_2
AN-63	66.00	sample	sample	CA22	mcontrol	E	control	5.221879068	0.1979625	46	57			21.45	6/6/2019	batch_11	seq_2
AN-74	20.00	sample	sample	CA23	mcontrol	U	control	2.6471469420	0.5215875	18	59			24.56	6/7/2019	batch_12	seq_1
AN-71	36.00	sample	sample	CA23	mcontrol	S	control	617.7574214	5.633285366	18	59			24.56	6/7/2019	batch_12	seq_1
AN-73	58.00	sample	sample	CA23	mcontrol	E	control	2.631906124	0.532375	18	59			24.56	6/7/2019	batch_12	seq_2
AN-72	66.00	sample	sample	CA23	mcontrol	S	control	11.77134109	1.783725	18	59			24.56	6/7/2019	batch_12	seq_1
AN-89	70.00	sample	sample	CA24	mcontrol	S	control	97.88561127	0.21731	30	62			22.77	6/20/2019	batch_14	seq_1
AN-92	76.00	sample	sample	CA24	mcontrol	U	control	3.362190314	7.07E-02	30	62			22.77	6/20/2019	batch_14	seq_2
AN-90	104.00	sample	sample	CA24	mcontrol	S	control	46.32998904	7.07E-02	30	62			22.77	6/20/2019	batch_14	seq_1
AN-91	262.00	sample	sample	CA24	mcontrol	E	control	18.25032552	0.43727	30	62			22.77	6/20/2019	batch_14	seq_1
AN-100	28.00	sample	sample	CA25	mcontrol	E	control	9.752135295	0.428105	46	71			22.92	6/21/2019	batch_15	seq_1
AN-99	38.00	sample	sample	CA25	mcontrol	S	control	1319.573538	1.821185	46	71			22.92	6/21/2019	batch_15	seq_1
AN-101	38.00	sample	sample	CA25	mcontrol	U	control	3.852691898	0.739715	46	71			22.92	6/21/2019	batch_15	seq_1
AN-111	15.40	sample	sample	CA26	mcontrol	E	control	2.0476136	0.34562	25	64.7			22.92	6/25/2019	batch_16	seq_2
AN-110	24.00	sample	sample	CA26	mcontrol	S	control	166.5586207	0.519755	25	64.7			22.92	6/25/2019	batch_16	seq_1
AN-112	28.00	sample	sample	CA26	mcontrol	U	control	1.04421065	0.739715	25	64.7			22.92	6/25/2019	batch_16	seq_1
AN-109	90.00	sample	sample	CA26	mcontrol	S	control	624.4890396	1.252955	25	64.7			22.92	6/25/2019	batch_16	seq_1
AN-291	52.00	sample	sample	CA27	mcontrol	S	control	91.75430772	0.54139	24	72			23.51	9/18/2020	batch_34	seq_2
AN-293	54.00	sample	sample	CA27	mcontrol	U	control	5.760258026	8.58E-02	24	72			23.51	9/18/2020	batch_34	seq_2
AN-292	60.00	sample	sample	CA27	mcontrol	E	control	3.163692866	0.985449	24	72			23.51	9/18/2020	batch_34	seq_2
AN-303	32.00	sample	sample	CA28	mcontrol	U	control	1.279008013	0.702866	24	54			19.13	9/18/2020	batch_34	seq_2
AN-302	34.00	sample	sample	CA28	mcontrol	E	control	1.416578906	0.472186	24	54			19.13	9/18/2020	batch_34	seq_2
AN-301	36.00	sample	sample	CA28	mcontrol	S	control	38.5933922	0.25304	24	54			19.13	9/18/2020	batch_34	seq_2
AN-311	32.00	sample	sample	CA29	mcontrol	S	control	139.6716838	0.2576055	16	62			24.21	9/22/2020	batch_35	seq_2
AN-313	32.00	sample	sample	CA29	mcontrol	U	control	1.559705047	0.525282	16	62			24.21	9/22/2020	batch_35	seq_2
AN-312	38.00	sample	sample	CA29	mcontrol	E	control	2.173933309	6.95E-02	16	62			24.21	9/22/2020	batch_35	seq_2
AN-176	26.00	sample	sample	CA3	mcontrol	U	control	2.37001069	1.7152375	19	52			19.33	9/12/2019	batch_23	seq_1
AN-175	30.00	sample	sample	CA3	mcontrol	E	control	3.744795	0.91915	19	52			19.33	9/12/2019	batch_23	seq_1
AN-174	50.00	sample	sample	CA3	mcontrol	S	control	71.7291106	1.6255375	19	52			19.33	9/12/2019	batch_23	seq_1
AN-316	300.00	sample	sample	CA30	mcontrol	E	control	15.38303098	0.279309	16	58			20.55	9/24/2020	batch_36	seq_2
AN-317	424.00	sample	sample	CA30	mcontrol	U	control	5.489506571	1.5308775	16	58			20.55	9/24/2020	batch_36	seq_2
AN-315	532.00	sample	sample	CA30	mcontrol	S	control	98.412198	0.1563225	16	58			20.55	9/24/2020	batch_36	seq_2
AN-325	112.00	sample	sample	CA31	mcontrol	S	control	364.6561984	0.854603	20	61			22.41	9/25/2020	batch_37	seq_2
AN-326	222.00	sample	sample	CA31	mcontrol	E	control	5.41171952	2.556435	20	61			22.41	9/25/2020	batch_37	seq_2
AN-327	570.00	sample	sample	CA31	mcontrol	U	control	167.6019291	1.260722	20	61			22.41	9/25/2020	batch_37	seq_2
AN-342	114.00	sample	sample	CA32	mcontrol	E	control	12.92519368	3.0882575	20	57			23.57	9/28/2020	batch_38	seq_2
AN-341	458.00	sample	sample	CA32	mcontrol	S	control	281.38371	2.0052735	20	57			23.57	9/28/2020	batch_38	seq_2
AN-343	1406.00	sample	sample	CA32	mcontrol	U	control	27.5043779	1.860231	20	57			23.57	9/28/2020	batch_38	seq_2
AN-347	198.00	sample	sample	CA33	mcontrol	U	control	2.547074634	0.448484	51	49			18.9	10/2/2020	batch_39	seq_2
AN-346	470.00	sample	sample	CA33	mcontrol	E	control	9.12606358	0.2454245	51	49			18.9	10/2/2020	batch_39	seq_2
AN-345	822.00	sample	sample	CA33	mcontrol	S	control	261.5033705	0.3227805	51	49			18.9	10/2/2020	batch_39	seq_2
AN-178	24.00	sample	sample	CA4	mcontrol	E	control	10.24064992	0.964	21	80			24.15	9/12/2019	batch_23	seq_1
AN-179	34.00	sample	sample	CA4	mcontrol	U	control	2.516102463	0.9752125	21	80			24.15	9/12/2019	batch_23	seq_1
AN-177	36.00	sample	sample	CA4	mcontrol	S	control	374.8	8.659915162	21	80			24.15	9/12/2019	batch_23	seq_1
AN-193	60.00	sample	sample	CA5	mcontrol	S	control	23.83988945	0.717325	18	61			22.41	9/23/2019	batch_23	seq_2
AN-194	30.00	sample	sample	CA5	mcontrol	E	control	3.796067543	0.65005	18	61			22.41	9/30/2019	batch_24	seq_2
AN-195	58.00	sample	sample	CA5	mcontrol	U	control	3.984562463	0.91915	18	61			22.41	9/30/2019	batch_24	seq_2
AN-197	20.00	sample	sample	CA6	mcontrol	E	control	4.03494252	1.390075	18	55			20.96	9/30/2019	batch_24	seq_2
AN-198	28.00	sample	sample	CA6	mcontrol	U	control	1.386293005	1.3340125	18	55			20.96	9/30/2019	batch_24	seq_1
AN-196	38.00	sample	sample	CA6	mcontrol	S	control	47.81654565	1.5919	18	55			20.96	9/30/2019	batch_24	seq_2
AN-213	40.00	sample	sample	CA7	mcontrol	E	control	17.39117551	1.0873375	22	65			23.88	10/7/2019	batch_26	seq_2
AN-214	46.00	sample	sample	CA7	mcontrol	U	control	7.213821025	2.1188875	22	65			23.88	10/7/2019	batch_26	seq_2
AN-212	54.00	sample	sample	CA7	mcontrol	S	control	181.2457465	0.91915	22	65			23.88	10/7/2019	batch_26	seq_1
AN-148	28.00	sample	sample	CA8	mcontrol	E	control	9.25387778	1.548634	24	54			18.69	8/15/2019	batch_20	seq_1
AN-149	30.00	sample	sample	CA8	mcontrol	U	control	2.087328927	2.519758	24	54			18.69	8/15/2019	batch_20	seq_1
AN-147	34.00	sample	sample	CA8	mcontrol	S	control	267.1033028	1.641122	24	54			18.69	8/15/2019	batch_20	seq_1
AN-166	32.00	sample	sample	CA9	mcontrol	U	control	0.793914993	0.493075	22	57			19.72	9/3/2019	batch_22	seq_1
AN-164	64.00	sample	sample	CA9	mcontrol	S	control	100.5103085	0.8406625	22	57			19.72	9/3/2019	batch_22	seq_1
AN-165	82.00	sample	sample	CA9	mcontrol	E	control	1.42695858	0.964	22	57			19.72	9/3/2019	batch_22	seq_2
Mock-1			mock	D0	mock	mock	mock									batch_0	seq_2
Mock-2			mock	D0	mock	mock	mock									batch_0	seq_2
Mock-3			mock	D0	mock	mock	mock									batch_0	seq_2
Mock-plate1			mock	D0	mock	mock	mock									batch_0	seq_1
Mock-plate2			mock	D0	mock	mock	mock									batch_0	seq_1
Mock-plate3			mock	D0	mock	mock	mock									batch_0	seq_1
AN-359	31080.00		mock	D1											10/7/2020	batch_40	seq_2
AN-360	15940.00		mock	D2											10/7/2020	batch_40	seq_2
AN-361	576.00		mock	D3											10/7/2020	batch_40	seq_2
AN-362	1320.00		mock	D4											10/7/2020	batch_40	seq_2
AN-363	622.00		mock	D5											10/7/2020	batch_40	seq_2
AN-364	284.00		mock	D6											10/7/2020	batch_40	seq_2
AN-365	130.00		mock	D7											10/7/2020	batch_40	seq_2
AN-366	36.00		mock	D8											10/7/2020	batch_40	seq_2
AN-368	16.40	control	kit_control	kit_control												batch_0	seq_2
AN-284	48.00	control	kit_control	kit_control											9/15/2020	batch_33	seq_2
AN-304	52.00	control	kit_control	kit_control											9/18/2020	batch_34	seq_2
AN-294	162.00	control	kit_control	kit_control											9/18/2020	batch_34	seq_2
AN-314	70.00	control	kit_control	kit_control											9/22/2020	batch_35	seq_2
AN-324	48.00	control	kit_control	kit_control											9/24/2020	batch_36	seq_2
AN-334	74.00	control	kit_control	kit_control											9/25/2020	batch_37	seq_2
AN-344	38.00	control	kit_control	kit_control											9/28/2020	batch_38	seq_2
AN-358	62.00	control	kit_control	kit_control											10/2/2020	batch_39	seq_2
AN-173	22.00	control	room_control	room_control											9/3/2019	batch_22	seq_1
AN-357	78.00	control	room_control	room_control											10/2/2020	batch_39	seq_2
AN-356	474.00	control	room_control	room_control											10/2/2020	batch_39	seq_2

Supplementary Table S2. Summarized AMP concentrations per group and sampling site		
Psoriasin first measurement	**AN subjects (n=33)**	**HC subjects (n=33)**
Forehead		
Mean (SD)	172.08 (155.48)	244.81 (258.81)
Median (IQR)	146.84 (195.37)	166.69 (230.22)
Range	1.64 - 664.65	23.84-1319.57
Elbow		
Mean (SD)	10.10 (9.92)	7.49 (4.93)
Median (IQR)	4.76 (13.9)	6.55 (6.21)
Range	0.80 - 34.81	1.42 - 18.25
Lower forearm		
Mean (SD)	5.98 (6.76)	8.91 (28.89)
Median (IQR)	2.71 5.22)	2.52 (3.80)
Range	0.90 - 27.23	0.73 - 167.60
Psoriasin second measurement	**AN subjects (n=32)**	**-**
Forehead		
Mean (SD)	273.93 (262.06)	-
Median (IQR)	208.88 (325.49)	-
Range	2.90 - 1144.75	
Elbow		
Mean (SD)	6.87 (5.54)	-
Median (IDR)s	4.81 (5.32)	-
Range	0.98 - 25.11	-
Lower forearm		
Mean (SD)	6.28 (5.20)	-
Median (IQR)	4.7 (6.27)	-
Range	0.6 - 17.39	-
Rnase 7 first measurement	**AN (n=33)**	**HC subjects (n =33)**
Forehead		
Mean (SD)	2.09 (4.16)	1.71 (1.81)
Median (IQR)	1.09 (1.10)	1.14 (1.40)
Range	0 - 23.52	0.07 - 8.66
Elbow		
Mean (SD)	1.16 (1.13)	1.17 (0.82)
Median (IQR)	0.8 (0.73)	0.99 (0.92)
Range	0.09 - 4.78	0.07 - 3. 70
Lower forearm		
Mean (SD)	1.36 (1.09)	1.28 (1.03)
Median (IQR)	0.95 (0.99)	0.98 (0.93)
Range	0.26 - 4.58	0.07 - 4.86
Rnase 7 second measurement	**AN (n=32)**	-
Forehead		
Mean (SD)	2.21 (2.97)	-
Median (IQR)	1.11 (1.32)	-
Range	0.16 - 14.31	-
Elbow		
Mean (SD)	1.21 (1.04)	-
Median (IQR)	0.82 (0.82)	-
Range	0.034 - 3.88	-
Lower forearm		
Mean (SD)	1.83 (2.71)	-
Median (IQR)	1.05 (1.7)	-
Range	0.05 - 15.16	-

Key: AMP = antimicrobial peptides; AN = anorexia nervoa; HC = healthy weight controls
SD = standard deviation; IQR = interquartile range

Supplementary Table S3. Summarized bacterial load data and associations with AMPs, BMI

	Mean (SD)	Median	Range
Healthy controls			
Sampling locations:			
Inner elbow (n = 30)	84.18 (110.21)	40	15.4-470.0
Forehead (n = 35)	163.26 (201.25)	60	24.0-822.0
Lower arm (n = 31)	128.71 (264.92)	48	20.0-1406.0
Total, all locations	127.4 (203.54)	50	15.4-1406.0
Patients with AN, before			
Sampling locations:			
Inner elbow (n = 30)	134.08 (231.90)	78	14.8-1176
Forehead (n = 35)	197.31 (255.89)	116	22.0-890.0
Lower arm (n = 31)	141.01 (337.14)	55	14.2-1840.0
Total, all locations	158.1 (279.08)	70	14.2-1840.0
Patients with AN, after			
Sampling locations:			
Inner elbow (n = 30)	322.59 (409.73)	120	16.6-1582.0
Forehead (n = 35)	330.35 (645.26)	94	2/17/3420
Lower arm (n = 31)	157.87 (219.92)	61	22.0-784.0
Total, all locations	267.7 (459.95)	78	16.6-3420.0

	AMP	Spearman's Rho	p-value | adj. p-value
Total, all locations	psoriasin	0.25	1.91E-05
Total, all locations	RNase 7	-0.04	0.55
Total, all locations	BMI	-0.07	0.26
Sampling locations:			
Inner elbow	psoriasin	0.19	0.075 | 0.23
Inner elbow	RNase 7	0.05	0.616 | 1.00
Inner elbow	BMI	-0.11	0.310 | 0.93
Forehead	psoriasin	0.28	0.006 | 0.02
Forehead	RNase 7	-0.13	0.214 | 0.64
Forehead	BMI	-0.05	0.630 | 1.00
Lower arm	psoriasin	0.22	0.034 | 0.10
Lower arm	RNase 7	-0.12	0.240 | 0.72
Lower arm	BMI	-0.49	0.644 | 1.00

AN = anorexia nervosa; before = before weight gain; after = after weight gain

Bacterial load was measured by ddPCR, as previously described by Sze et al. (2014)

p -values were adjusted for multiple testing according to Benjamini and Hochberg (1995)

Supplementary Table S4. Summarized statistical analyses for mean relative abundances of indicator taxa at the phylum and genus levels

Proteobacteria phylum							
	Group 1	**Group 2**	**Test**	**Test statistic**	***n* samples**	**Adjusted *p*-value***	**Signifiicance level**
	AN before	HC	Wilcoxon, unpaired	W = 5055.5	191	0.014	*
	AN before	AN after	Wilcoxon, paired	V = 2584	93	0.005	**
	HC	AN after	Wilcoxon, unpaired	W = 4055	191	0.736	ns
Firmicutes phylum							
	Group 1	**Group 2**	**Test**	**Test statistic**	***n* samples**	**Adjusted *p*-value***	**Signifiicance level**
	AN before	HC	Wilcoxon, unpaired	W = 3119	191	0.003	**
	AN before	AN after	Wilcoxon, paired	V = 1478.5	93	0.066	ns
	HC	AN after	Wilcoxon, unpaired	W = 4056	191	0.738	ns
Actinobacteria phylum							
	Group 1	**Group 2**	**Test**	**Test statistic**	***n* samples**	**Adjusted *p*-value***	**Signifiicance level**
	AN before	HC	Wilcoxon, unpaired	W = 4228	191	0.886	ns
	AN before	AN after	Wilcoxon, paired	V = 1992	93	0.743	ns
	HC	AN after	Wilcoxon, unpaired	W = 4129.5	191	0.898	ns
Bacteroidetes phylum							
	Group 1	**Group 2**	**Test**	**Test statistic**	***n* samples**	**Adjusted *p*-value***	**Signifiicance level**
	AN before	HC	Wilcoxon, unpaired	W = 3880	191	0.407	ns
	AN before	AN after	Wilcoxon, paired	V = 1850	93	0.271	ns
	HC	AN after	Wilcoxon, unpaired	W = 3510	191	0.062	ns
Lactobacillus genus							
	Group 1	**Group 2**	**Test**	**Test statistic**	***n* samples**	**Adjusted *p*-value***	**Signifiicance level**
	AN before	HC	Wilcoxon, unpaired	W = 2462	191	4.23E-07	****
	AN before	AN after	Wilcoxon, paired	V = 762	93	0.51	ns
	HC	AN after	Wilcoxon, unpaired	W = 2326.5	191	5.12E-08	****
Staphyloccocus genus							
	Group 1	**Group 2**	**Test**	**Test statistic**	***n* samples**	**Adjusted *p*-value***	**Signifiicance level**
	AN before	HC	Wilcoxon, unpaired	W = 3350.5	191	0.021	*
	AN before	AN after	Wilcoxon, paired	V = 1213	93	0.005	**
	HC	AN after	Wilcoxon, unpaired	W = 4747.5	191	0.111	ns
Anaerococcus							
	Group 1	**Group 2**	**Test**	**Test statistic**	***n* samples**	**Adjusted *p*-value***	**Signifiicance level**
	AN before	HC	Wilcoxon, unpaired	W = 3523.5	191	0.0620	ns
	AN before	AN after	Wilcoxon, paired	V = 1222	93	0.9330	ns
	HC	AN after	Wilcoxon, unpaired	W = 3684.5	191	0.1630	ns
Propionibacterium genus							
	Group 1	**Group 2**	**Test**	**Test statistic**	***n* samples**	**Adjusted *p*-value***	**Signifiicance level**
	AN before	HC	Wilcoxon, unpaired	W = 4342.5	191	0.6240	ns
	AN before	AN after	Wilcoxon, paired	V = 1118.5	93	0.6020	ns
	HC	AN after	Wilcoxon, unpaired	W = 3968.5	191	0.5	ns
Streptococcus genus							
	Group 1	**Group 2**	**Test**	**Test statistic**	***n* samples**	**Adjusted *p*-value***	**Signifiicance level**
	AN before	HC	Wilcoxon, unpaired	W = 4327.5	191	0.582	ns
	AN before	AN after	Wilcoxon, paired	V = 1840	93	0.292	ns
	HC	AN after	Wilcoxon, unpaired	W = 4221.5	191	0.922	ns
Micrococcus genus							
	Group 1	**Group 2**	**Test**	**Test statistic**	***n* samples**	**Adjusted *p*-value***	**Signifiicance level**
	AN before	HC	Wilcoxon, unpaired	W = 4206.5	191	0.921	ns
	AN before	AN after	Wilcoxon, paired	V = 726.5	93	0.259	ns
	HC	AN after	Wilcoxon, unpaired	W = 4153	191	0.941	ns
Corynebacterium genus							
	Group 1	**Group 2**	**Test**	**Test statistic**	***n* samples**	**Adjusted *p*-value***	**Signifiicance level**
	AN before	HC	Wilcoxon, unpaired	W = 3835	191	0.340	ns
	AN before	AN after	Wilcoxon, paired	V = 1435	93	0.376	ns
	HC	AN after	Wilcoxon, unpaired	W = 47403488	191	0.692	ns
Acinetobacter							
	Group 1	**Group 2**	**Test**	**Test statistic**	***n* samples**	**Adjusted *p*-value***	**Signifiicance level**
	AN before	HC	Wilcoxon, unpaired	W = 4228	191	0.886	ns
	AN before	AN after	Wilcoxon, paired	V = 1992	93	0.743	ns
	HC	AN after	Wilcoxon, unpaired	W = 4129.5	191	0.898	ns
unclassified Streptophyta							
	Group 1	**Group 2**	**Test**	**Test statistic**	***n* samples**	**Adjusted *p*-value***	**Signifiicance level**
	AN before	HC	Wilcoxon, unpaired	W = 4712	191	0.1030	ns
	AN before	AN after	Wilcoxon, paired	V = 1135.5	93	0.9850000	ns
	HC	AN after	Wilcoxon, unpaired	W = 4826.5	191	0.0490	*
unclassified Neisseriaceae							
	Group 1	**Group 2**	**Test**	**Test statistic**	***n* samples**	**Adjusted *p*-value***	**Signifiicance level**
	AN before	HC	Wilcoxon, unpaired	W = 4918.5	191	0.003	**
	AN before	AN after	Wilcoxon, paired	V = 519	93	0.402	ns
	HC	AN after	Wilcoxon, unpaired	W = 4717	191	0.026	*

Mean relative abundance *Lactobacillus* per sampling location							
Location	**Group 1**	**Group 2**	**Test**			**Adjusted *p* -value***	**Signifiicance level**
Forehead	AN before	HC	Wilcoxon, unpaired	W = 733.5	191	0.001	***
Forehead	AN before	AN after	Wilcoxon, paired	V = 68	93	0.346	ns
Forehead	HC	AN after	Wilcoxon, unpaired	W = 781.5	191	9.83E-05	****
Elbow	AN before	HC	Wilcoxon, unpaired	W = 568.5	191	0.015	*
Elbow	AN before	AN after	Wilcoxon, paired	V = 77.5	93	0.642	ns
Elbow	HC	AN after	Wilcoxon, unpaired	W = 594	191	0.004	**
Lower forearm	AN before	HC	Wilcoxon, unpaired	W = 671	191	0.002	**
Lower forearm	AN before	AN after	Wilcoxon, paired	V = 77.5	93	0.728	ns
Lower forearm	HC	AN after	Wilcoxon, unpaired	W = 661.5	191	0.003	**
Mean relative abundance *Abiotrophia* per sampling location							
Location	**Group 1**	**Group 2**	**Test**			**Adjusted *p* -value***	**Signifiicance level**
Forehead	AN before	HC	Wilcoxon, unpaired	W = 603.5	191	0.067	ns
Forehead	AN before	AN after	Wilcoxon, paired	V = 9.8	93	0.68	ns
Forehead	HC	AN after	Wilcoxon, unpaired	W = 621	191	0.03	*
Elbow	AN before	HC	Wilcoxon, unpaired	W = 463	191	0.174	ns
Elbow	AN before	AN after	Wilcoxon, paired	V = 1.5	93	1	ns
Elbow	HC	AN after	Wilcoxon, unpaired	W = 463	191	0.174	ns
Lower forearm	AN before	HC	Wilcoxon, unpaired	W = 479	191	0.682	ns
Lower forearm	AN before	AN after	Wilcoxon, paired	V = 3	93	0.37	ns
Lower forearm	HC	AN after	Wilcoxon, unpaired	W = 494.5	191	0.33	ns
Mean relative abundance *Clostridium* per sampling location							
Location	**Group 1**	**Group 2**	**Test**			**Adjusted *p* -value***	**Signifiicance level**
Forehead	AN before	HC	Wilcoxon, unpaired	W = 594.5	191	0.021	*
Forehead	AN before	AN after	Wilcoxon, paired	NA	NA	NA	NA
Forehead	HC	AN after	Wilcoxon, unpaired	W = 594.5	191	0.021	*
Elbow	AN before	HC	Wilcoxon, unpaired	W = 450.5	191	0.478	ns
Elbow	AN before	AN after	Wilcoxon, paired	V = 16.5	93	0.24	ns
Elbow	HC	AN after	Wilcoxon, unpaired	W = 480	191	0.122	ns
Lower forearm	AN before	HC	Wilcoxon, unpaired	W = 484	191	0.665	ns
Lower forearm	AN before	AN after	Wilcoxon, paired	V = 10	93	0.1	ns
Lower forearm	HC	AN after	Wilcoxon, unpaired	W = 540	191	0.024	*
Mean relative abundance unlcassified Neisseriaceae per sampling location							
Location	**Group 1**	**Group 2**	**Test**			**Adjusted *p* -value***	**Signifiicance level**
Forehead	AN before	HC	Wilcoxon, unpaired	W = 369	191	0.026	*
Forehead	AN before	AN after	Wilcoxon, paired	V = 136	93	0.254	ns
Forehead	HC	AN after	Wilcoxon, unpaired	W = 404.5	191	0.09	ns
Elbow	AN before	HC	Wilcoxon, unpaired	W = 329.5	191	0.015	*
Elbow	AN before	AN after	Wilcoxon, paired	V = 45.5	93	0.286	ns
Elbow	HC	AN after	Wilcoxon, unpaired	W = 374	191	0.146	ns
Lower forearm	AN before	HC	Wilcoxon, unpaired	W = 420	191	0.298	ns
Lower forearm	AN before	AN after	Wilcoxon, paired	V = 29	93	0.456	ns
Lower forearm	HC	AN after	Wilcoxon, unpaired	W = 6415	191	0.247	ns
Mean relative abundance Staphylococcus per sampling location							
Location	**Group 1**	**Group 2**	**Test**				
Forehead	AN before	HC	Wilcoxon, unpaired	W = 631	191	0.100	ns
Forehead	AN before	AN after	Wilcoxon, paired	V = 141	93	0.100	ns
Forehead	HC	AN after	Wilcoxon, unpaired	W = 465.5	191	0.576	ns
Elbow	AN before	HC	Wilcoxon, unpaired	W = 506	191	0.180	ns
Elbow	AN before	AN after	Wilcoxon, paired	V = 113	93	0.040	ns
Elbow	HC	AN after	Wilcoxon, unpaired	W = 349.5	191	0.276	ns
Lower forearm	AN before	HC	Wilcoxon, unpaired	W = 521	191	0.423	ns
Lower forearm	AN before	AN after	Wilcoxon, paired	V = 165	93	0.171	ns
Lower forearm	HC	AN after	Wilcoxon, unpaired	W = 379	191	0.217	ns
Mean relative abundance Propionibacterium per sampling location							
Location	**Group 1**	**Group 2**	**Test**				
Forehead	AN before	HC	Wilcoxon, unpaired	W = 453.5	191	0.467	ns
Forehead	AN before	AN after	Wilcoxon, paired	V = 207	93	0.674	ns
Forehead	HC	AN after	Wilcoxon, unpaired	W = 508.5	191	0.995	ns
Elbow	AN before	HC	Wilcoxon, unpaired	W = 477	191	0.307	ns
Elbow	AN before	AN after	Wilcoxon, paired	V = 51	93	0.63	ns
Elbow	HC	AN after	Wilcoxon, unpaired	W = 450	191	0.603	ns
Lower forearm	AN before	HC	Wilcoxon, unpaired	W = 393.5	191	0.264	ns
Lower forearm	AN before	AN after	Wilcoxon, paired	V = 119	93	0.344	ns
Lower forearm	HC	AN after	Wilcoxon, unpaired	W = 486	191	0.73	ns

p -values: * <0.05; ** < 0.01; *** < 0.001; **** < 0.0001

Key: ns = not significant; HC = healthy-weight control; AN before = anorexia nervosa before weight gain; AN after = anorexia nervosa after weight gain

p -values were adjusted for multiple testing according to Benjamini and Hochberg (1995).

Supplementary Table S5. Alpha diversity statistics

Shannon diversity by sampling location

Location	Group 1	Group 2	Test			Adjusted *p* -value*	Signifiicance level
Forehead	AN before	HC	Wilcoxon, unpaired	W = 617	191	0.142	ns
Forehead	AN before	AN after	Wilcoxon, paired	V = 460	93	0.547	ns
Forehead	HC	AN after	Wilcoxon, unpaired	W = 682	191	0.02	*
Elbow	AN before	HC	Wilcoxon, unpaired	W = 403	191	0.799	ns
Elbow	AN before	AN after	Wilcoxon, paired	V = 493	93	0.1	ns
Elbow	HC	AN after	Wilcoxon, unpaired	W = 519	191	0.126	ns
Lower forearm	AN before	HC	Wilcoxon, unpaired	W = 571	191	0.129	ns
Lower forearm	AN before	AN after	Wilcoxon, paired	V = 487	93	0.592	ns
Lower forearm	HC	AN after	Wilcoxon, unpaired	W = 659	191	0.005	**

Chao1 index by sampling location

Location	Group 1	Group 2	Test			Adjusted *p* -value*	Signifiicance level
Forehead	AN before	HC	Wilcoxon, unpaired	W = 682	191	0.02	*
Forehead	AN before	AN after	Wilcoxon, paired	V = 408	93	0.852	ns
Forehead	HC	AN after	Wilcoxon, unpaired	W = 708.5	191	0.007	**
Elbow	AN before	HC	Wilcoxon, unpaired	W = 390	191	0.646	ns
Elbow	AN before	AN after	Wilcoxon, paired	V = 456.5	93	0.294	ns
Elbow	HC	AN after	Wilcoxon, unpaired	W = 465	191	0.488	ns
Lower forearm	AN before	HC	Wilcoxon, unpaired	W = 585	191	0.085	ns
Lower forearm	AN before	AN after	Wilcoxon, paired	V = 486	93	0.599	ns
Lower forearm	HC	AN after	Wilcoxon, unpaired	W = 646.5	191	0.009	**

p -values: * <0.05; ** < 0.01; *** < 0.001; **** < 0.0001

Key: ns = not significant; HC = healthy-weight control; AN before = anorexia nervosa before weight gain; AN after = anorexia nervosa after weight gain

p -values were adjusted for multiple testing according to Benjamini and Hochberg (1995).

Supplementary Table S6. Spearman's correlations between alpha diversity measures and BMI, bacterial load, and AMP concentrations

Alpha diversity index	Feature	rho (spearman's r)	adj. *p* -value	Significance level
Chao1 index	BMI	0.104	0.088	ns
Shannon's diversity index	BMI	0.056	0.36	ns
Chao1 index	total bacterial load	-0.041	0.5	ns
Shannon's diversity index	total bacterial load	-0.35	4.60E-09	****
At sampling locations:				
Forehead	total bacterial load	-0.033	0.003	**
Inner elbow	total bacterial load	-0.350	0.003	**
Lower arm	total bacterial load	-0.260	0.04	*
Chao1 index	Rnase 7	0.040	0.50	ns
Shannon's diversity index	Rnase 7	-0.050	0.42	ns
Chao1 index	psoriasin	0.007	0.91	ns
Shannon's diversity index	psoriasin	-0.220	0.00028	***
At sampling locations:				
Forehead	psoriasin	-0.055	1.0	ns
Inner elbow	psoriasin	-0.063	1.0	ns
Lower arm	psoriasin	-0.098	0.36	ns

p -values: * <0.05; ** < 0.01; *** < 0.001; **** < 0.0001

ns = not significant

p -values were adjusted for multiple testing according to Benjamini and Hochberg (1995).

Supplementary Table S7. Spearman's correlations between relative abundances of indicator genera and AMP concentrations and BMI, per sampling site

Forehead					
Genus	**Feature**	**Test**	**rho (spearman's r)**	**adj. *p* -value**	**Significance level**
Lactobacillus	Psoriasin	Spearman's	-0.06	0.56	ns
Lactobacillus	RNase 7	Spearman's	0.16	0.14	ns
Lactobacillus	BMI	Spearman's	0.49	1.65E-09	****
Abiotrophia	Psoriasin	Spearman's	0.16	0.12	ns
Abiotrophia	RNase 7	Spearman's	-0.14	0.67	ns
Abiotrophia	BMI	Spearman's	0.15	0.16	ns
Clostridium	Psoriasin	Spearman's	0.14	0.18	ns
Clostridium	RNase 7	Spearman's	-0.2	0.157	ns
Clostridium	BMI	Spearman's	0.27	0.024	*
Inner elbow					
Genus	**Feature**	**Test**	**rho (spearman's r)**	**adj. *p* -value**	**Significance level**
Lactobacillus	Psoriasin	Spearman's	-0.044	0.69	ns
Lactobacillus	RNase 7	Spearman's	0.1	0.37	ns
Lactobacillus	BMI	Spearman's	0.3	0.0054	**
Abiotrophia	Psoriasin	Spearman's	-0.0061	0.96	ns
Abiotrophia	RNase 7	Spearman's	0.086	0.43	ns
Abiotrophia	BMI	Spearman's	0.12	0.27	ns
Clostridium	Psoriasin	Spearman's	0.13	0.24	ns
Clostridium	RNase 7	Spearman's	0.04	1	ns
Clostridium	BMI	Spearman's	0.77	0.5	ns
Lower arm					
Genus	**Feature**	**Test**	**rho (spearman's r)**	**adj. *p* -value**	**Significance level**
Lactobacillus	Psoriasin	Spearman's	0.01	0.93	ns
Lactobacillus	RNase 7	Spearman's	0.12	0.27	ns
Lactobacillus	BMI	Spearman's	0.34	0.0014	**
Abiotrophia	Psoriasin	Spearman's	0.14	0.2	ns
Abiotrophia	RNase 7	Spearman's	0.15	0.48	ns
Abiotrophia	BMI	Spearman's	0.04	0.72	ns
Clostridium	Psoriasin	Spearman's	0.17	0.12	ns
Clostridium	RNase 7	Spearman's	-0.082	1	ns
Clostridium	BMI	Spearman's	0.04	0.7	ns
Genus	**Feature**	**Test**	**rho (spearman's r)**	***p* -value**	**Significance level**
Lactobacillus	bacterial load	Spearman's	-0.051	0.41	ns
Abiotrophia	bacterial load	Spearman's	0.009	0.89	ns
Clostridium	bacterial load	Spearman's	-0.061	0.31	ns
Uncl. *Neisseriaceae*	bacterial load	Spearman's	0.043	0.48	ns

p -values: * <0.05; ** < 0.01; *** < 0.001; **** < 0.0001

ns = not significant

AMP = anti-microbial peptide; ASV = amplicon-sequence variant

p -values were adjusted for multiple testing according to Benjamini and Hochberg (1995).

Supplementary Table S8. RDP SeqMatch results for indicator species at the ASV-level

ASV_29

Sequence ID	Similarity score	S_ab score	unique common oligomers	sequence full name
S000126968	**not_calculated**	**1.000**	**0689**	**Lactobacillus crispatus; 180; AJ421225**
S000142938	**not_calculated**	**1.000**	**0634**	**Lactobacillus crispatus; BJ H42-4; AY339181**
S000142939	**not_calculated**	**1.000**	**0617**	**Lactobacillus crispatus; BJ Y20; AY339182**
S000143207	not_calculated	1.000	0625	Lactobacillus gasseri; BJ H36-3b; AY339179
S000368361	not_calculated	1.000	0871	Lactobacillus crispatus; FX13-3; AY335493
S000368363	not_calculated	1.000	0871	Lactobacillus crispatus; FX36-3; AY335495
S000389900	not_calculated	1.000	1413	Lactobacillus crispatus; KC35b; AF243158
S000389914	not_calculated	1.000	1410	Lactobacillus crispatus; TL25a; AF243172
S000390171	not_calculated	1.000	1424	Lactobacillus crispatus (T); ATCC33820; AF257097
S000400895	not_calculated	1.000	0634	Lactobacillus crispatus; TSK V36-1; AY190621
S000435488	not_calculated	1.000	0642	Lactobacillus crispatus; TSK V38-1; AY190616
S000536099	not_calculated	1.000	1418	uncultured bacterium; rRNA022; AY958795
S000536148	not_calculated	1.000	1768	uncultured bacterium; rRNA071; AY958844
S000536176	not_calculated	1.000	1418	uncultured bacterium; rRNA099; AY958872
S000536190	not_calculated	1.000	1444	uncultured bacterium; rRNA113; AY958886
S000536204	not_calculated	1.000	1410	uncultured bacterium; rRNA127; AY958900
S000536207	not_calculated	1.000	1442	uncultured bacterium; rRNA130; AY958903
S000536233	not_calculated	1.000	1446	uncultured bacterium; rRNA156; AY958929
S000536281	not_calculated	1.000	1392	uncultured bacterium; rRNA204; AY958977
S000536310	not_calculated	1.000	1408	uncultured bacterium; rRNA233; AY959006

ASV_160

Sequence ID	Similarity score	S_ab score	unique common oligomers	sequence full name
S001172347	not_calculated	1.000	1446	uncultured Abiotrophia sp.; 2.1; DQ346440
S001889319	**not_calculated**	**1.000**	**0466**	**Abiotrophia defectiva; IL025; GU411212**
S001889348	**not_calculated**	**1.000**	**0454**	**Abiotrophia defectiva; RN041; GU411241**
S002083844	not_calculated	1.000	1308	uncultured bacterium; ncd563g04c1; HM278818
S002110434	not_calculated	1.000	1304	uncultured bacterium; ncd850f04c1; HM305408
S002113776	not_calculated	1.000	1306	uncultured bacterium; ncd902c05c1; HM308750
S002118822	not_calculated	1.000	1301	uncultured bacterium; ncd385h07c1; HM313796
S002118885	not_calculated	1.000	1303	uncultured bacterium; ncd388a06c1; HM313859
S002126630	not_calculated	1.000	1298	uncultured bacterium; ncd385a05c1; HM321604
S002128420	not_calculated	1.000	1303	uncultured bacterium; ncd411h09c1; HM323394
S002128440	not_calculated	1.000	1301	uncultured bacterium; ncd412c05c1; HM323414
S002609210	not_calculated	1.000	1305	uncultured bacterium; ncd1282a02c1; JF089277
S002628349	not_calculated	1.000	1305	uncultured bacterium; ncd1327e06c1; JF108416
S002738922	not_calculated	1.000	1304	uncultured bacterium; ncd2568a04c1; JF218989
S002745609	not_calculated	1.000	1304	uncultured bacterium; ncd2566a11c1; JF225676
S002747042	not_calculated	1.000	1302	uncultured bacterium; ncd2590e01c1; JF227109
S002747485	not_calculated	1.000	1302	uncultured bacterium; ncd2597d05c1; JF227552
S002747629	not_calculated	1.000	1303	uncultured bacterium; ncd2600a01c1; JF227696
S002747700	not_calculated	1.000	1302	uncultured bacterium; ncd2600f08c1; JF227767
S002749638	not_calculated	1.000	1301	uncultured bacterium; ncd2630g02c1; JF229705

ASV_744

Sequence ID	Similarity score	S_ab score	unique common oligomers	sequence full name
S000388866	not_calculated	0.872	1406	Clostridium sp. 45; AF191251
S000391440	not_calculated	0.872	1366	Clostridium thiosulfatireducens; Lup21; AF317650
S000395512	**not_calculated**	**1.000**	**0777**	**Clostridium sp. V13; AF502398**
S000434520	not_calculated	0.872	1367	Clostridium thiosulfatireducens; LUP 21; AY024332
S000603875	not_calculated	0.879	0747	uncultured Clostridium sp.; F1-11; DQ178973
S000705303	not_calculated	0.952	1405	iron-reducing enrichment clone Cl-W3; DQ677016
S000804481	not_calculated	0.872	1286	Clostridium thiosulfatireducens; DSM 13105; AB294141
S000805548	not_calculated	0.872	1420	Clostridium sulfidigenes (T); SGB2; EF199998
S000827154	not_calculated	0.872	1346	uncultured bacterium; lcfa_Bc84; DQ339700
S001169692	not_calculated	0.872	1160	Clostridium thiosulfatireducens; MG-2; EU937735
S002198190	not_calculated	0.875	1388	uncultured soil bacterium; D2B120; HM131959
S002448272	not_calculated	0.882	0796	Clostridium sp. LKS-AN-7; JF502819
S002747609	not_calculated	0.934	1265	uncultured bacterium; ncd2599b09c1; JF227676
S002948624	not_calculated	0.907	1409	uncultured bacterium; CA_88; JN559538
S002948670	not_calculated	0.893	1414	uncultured bacterium; CA_142; JN559584
S002948791	not_calculated	0.879	1409	uncultured bacterium; CA_277; JN559705
S003385735	not_calculated	0.934	0916	uncultured bacterium; WLCLC410; JN168389
S003931029	not_calculated	0.889	1350	Clostridium sp. A1; JN688046
S004010289	not_calculated	0.889	1350	bacterium enrichment culture clone M06; JN688024
S004092086	not_calculated	0.965	1391	Clostridium sp. Nesulana2; KJ722510

Seqmatch:version 3
RDP Data:release11_6
Data Set:both type and non-type strains,
:both environmental (uncultured) sequences and isolates,
:near-full-length sequences (>=1200 bases),
:good quality sequences
Comments:1558788 sequences were included in the search. The screening was based on 7-base oligomers.
Query Submit Date:Thu May 27 10:30:13 EDT 2021
Note:Orientation "-" means the query sequence has been reverse-complemented when the match is performed

Chapter 3

Characterization of skin microbiota in a murine *Nell2* knockout model

Hermes, B. M.[1,2,3], Hirose, M.[2], Tietje, A. M.[2], Belheouane, M.[1,4], Ibrahim, S.[2,5]*, Baines, J. F.[1,3]*
Characterization of skin microbiota in a murine *Nell2* knockout model. Manuscript in preparation.

Affiliations
[1]Section of Evolutionary Medicine, Institute for Experimental Medicine, Kiel University, Arnold-Heller-Str. 3, 24105 Kiel, Germany
[2]Lübeck Institute of Experimental Dermatology (LIED), University of Lübeck, Lübeck, Germany
[3]Max Planck Institute for Evolutionary Biology, August-Thienemann-Str. 2, 24306 Plön, Germany
[4]Research Center Borstel, Evolution of the Resistome, Leibniz Lung Center, Parkallee 1-40, 23845 Borstel, Germany
[5]College of Medicine and Sharjah Institute for Medical Research, University of Sharjah, 27272 Sharjah, UAE

*These authors contributed equally to this work.

Abstract
Skin microbiota play a crucial role in skin biology, including moderating local inflammatory responses and immune cell functioning. Disruptions in the homeostasis between host and commensal skin microbiota may lead to chronic inflammatory skin diseases. Thus, characterizing the relationship between host genetics and the assembly of the skin microbiome is central to understanding how microbiota influence human health and whether microbiota could be exploited as therapeutic interventions. Previously, using the 15th generation of an advanced intercross line, we demonstrated that abundances of bacterial taxa in the skin might be significantly influenced by host genetic variation. One exceptional candidate region was associated with unclassified *Betaproteobacteria* and contained one gene: neural epidermal growth factor-like 2 (*Nell2*). *Nell2* is predominately expressed in neural tissues but has also been found to be differentially expressed in the epidermis of patients suffering from atopic dermatitis (AD). While the relationship between *Nell2* and AD has not yet been elucidated, it is intriguing that an increased number of cutaneous free nerve endings has been observed in the epidermis of patients with AD, perhaps contributing to the intense pruritis that epitomizes this inflammatory skin disease.

Here, we aimed to further explore the association between *Nell2* and the associated taxon, unclassified *Betaproteobacteria*, in more detail through the analysis of a *Nell2* knock-out strain and by more precisely identifying the bacterial taxon involved through 16S rRNA gene amplicon sequencing. We reveal evidence suggesting that the unclassified *Betaproteobacteria* trait might instead belong to *Burkholderiaceae* within the class *Gammaproteobacteria*. Moreover, we find that unclassified *Betaproteobacteria* abundance does not significantly vary according to the examined *Nell2* genotype in the knock-out strain. We show that features of the skin microbiota do not significantly differ between *Nell2* genotypes. Finally, we find evidence suggesting that the unclassified *Betaproteobacteria* trait might be a common contaminant frequently found in DNA/RNA extraction kits. Our findings warrant future studies aimed at validating host gene-microbe associations previously observed in genetic mapping studies involving murine skin.

Background
Skin microbiota play a crucial role in several aspects of skin biology, including protective

immunity through the control of local inflammatory responses and immune cell functioning[1]. Recent studies have shown that host genotype contributes to skin microbiota variation[2,3], while bacterial traits may contribute to non-infectious skin pathologies, including psoriasis[4], acne[4], atopic dermatitis[1], cancer[2], and autoimmune diseases[3]. Disturbances in the homeostatic interplay between host and commensal skin microbiota may lead to chronic inflammatory diseases afflicting the skin. Characterizing the relationship between host genetics and the assembly of the skin microbiome is central to understanding how microbiota influence human health and whether microbiota could be exploited as therapeutic interventions for disease[5,6].

Using the 15th generation (G15) of an advanced intercross line (AIL), we previously demonstrated that, like gut microbiota, abundances of bacterial taxa in the skin might be significantly influenced by host genetic variation[2]. Remarkably, the combination of highly recombined individuals and 53,203 informative SNPs allowed the identification of genomic intervals as small as <0.1 megabases. The identified genomic intervals identified contained genes involved in skin inflammation and cancer and in some cases have known genotoxic or probiotic capabilities. One candidate region was associated with unclassified *Betaproteobacteria* and contained only one gene: neural epidermal growth factor-like 2 (*Nell2*).

Nell2 is expressed in neural tissues and has been reported to play a central role in the proliferation and differentiation of neural cells[7]. In 2010, Kamsteeg and colleagues[8] found *Nell2* to be differentially expressed in the epidermis of patients suffering from atopic dermatitis. A follow-up study utilizing a human skin equivalent exhibiting atopic dermatitis characteristics demonstrated an increase in *Nell2* expression after stimulation with Th2 cytokines, which are associated with atopy, but not with the stimulation of Th17-related or psoriasis-related cytokines[9,10]. While the relationship between *Nell2* and AD has not yet been elucidated, it is intriguing that Urashima and colleagues[11] reported an increased number of cutaneous free nerve endings in the epidermis of patients with AD, perhaps contributing to the intense pruritis that epitomizes atopic dermatitis.

Here, we aimed to further explore the association between *Nell2* and the associated taxon, unclassified *Betaproteobacteria*, in more detail through the analysis of a *Nell2* knock-out strain and by more precisely identifying the bacterial taxon involved through 16S rRNA gene amplicon sequencing. Accordingly, we constructed 16S rRNA partial gene clone libraries using a representative OTU sequence for the unclassified *Betaproteobacteria* trait and cDNA samples previously used in the Belheouane mapping study to identify bacterial candidates.
Samples and data derived from Belheouane and colleagues[2] are hereafter referred to as "G15 AIL samples" or the "G15 AIL dataset." Then, using a *Nell2* knock-out strain, we characterized the community composition of the skin microbiota according to *Nell2* genotype and searched for correlations between skin microbiota and any identified bacterial candidates. Samples and data derived from the *Nell2* knock-out strain are hereafter referred to as "*Nell2* mice samples" or the "*Nell2* mice dataset."

Results

Bacterial candidate identification—G15 AIL samples

First, we set out to precisely identify the candidate taxon from class *Betaproteobacteria* associated with *Nell2*, as reported by Belheouane and colleagues[2]. We thus aimed to amplify, clone, and sequence a longer portion of the 16S rRNA gene from the candidate taxon. For this, microbial candidate OTU_00001 was selected as the representative 16S rRNA gene sequence for the unclassified *Betaproteobacteria* trait, which represents the most prevalent unclassified *Betaproteobacteria* sequence in the G15 AIL dataset. Additionally, a subset of ten cDNA samples containing a high prevalence of unclassified *Betaproteobacteria* were selected from the G15 AIL dataset for primer testing and PCR amplification. These ten cDNA samples were previously used for QTL mapping by Belheouane and colleagues[2].

We first queried sequence OTU_00001 in RDP SeqMatch (v.3, release 11_6)[12] to gain insight into the taxonomy of the candidate trait. Results based on the highest SeqMatch score (S_ab) suggest OTU_00001 belongs to *Microvirgula*, from the family *Neisseriaceae* and class *Betaproteobacteria*. Thereafter, a literature search was conducted in the NCBI database for published primer pairs that amplify the V1-V2 hypervariable portion of the 16S rRNA gene. Primer pairs that target *Betaproteobacteria* at the class-level were selected for PCR testing (Supplementary Table 1). Additionally, a genus-specific primer pair for *Microvirgula* was designed using the NCBI primer-BLAST tool Primer3[13]. Query results in the NCBI primer design database indicated high specificity for *Microvirgula* species, but not related taxa. Two sets of primer pairs successfully amplified the V1-V2 hypervariable portion of the 16S rRNA gene and were selected to construct clone libraries (Table 1).

Table 1. Primer pairs amplifying the V1-V2 hypervariable portion of the target 16s rRNA gene

Primer name	**Pub. name**	**Sequence (5'-3')**	**Taxon**	**Ref.**
F2_micro	n/a	GGGGAGCTTGCTCCYGCTGA	Microvirgula	This paper
R7_micro	n/a	ACCCTACCCACTTCTGGCGGATTC	Microvirgula	This paper
Beta12_forw.	27F	AGAGTTTGATCMTGGCTCAG	Betaproteobact.	Ref. [14]
Beta12_reverse	682R	ACGCATTTCACTGCTACACG	Betaproteobact.	Ref. [15]

The first pair consisted of the novel primer pair F2_micro and R7_micro, which successfully amplified samples and a *Microvirgula* type strain (see Methods; Supplementary Figure 1). Subsequent sanger sequencing was successful for six of the 24 clone colonies. Taxonomy of aligned clone sequences was however indeterminate; RDP query results suggest clone sequences most closely match fungal genera (Supplementary Table 3).

The second primer pair, Beta12_forward and Beta12_reverse (hereafter, "Beta12") successfully amplified samples and bacteria type strains from *Neisseriaceae*, a family within the *Betaproteobacteria* class (see Methods; Supplementary Figure 2). Sanger sequencing was successful for 18 of the 24 clone colonies. Sequence analysis produced 13 unique sequences after removal of chimeric sequences. To refine taxonomic classification, clone sequences were queried using RDP SeqMatch[16]. Results suggest the clones best match with unclassified *Neisseriaceae* (S_ab scores range from 0.83 to 0.88) (Supplementary Table 4). As an additional means to identify similar 16S rRNA gene sequences, aligned clone sequences were queried using NCBI's basic local alignment search tool (MegaBLAST)[17]. The highest identity hits are based on percent identity and grade, which is a calculated percentage score by Geneious Prime[18] that combines percent identity, the e-value, and query coverage. The highest identity results comprise many genera from the family *Neisseriaceae* including *Eikenella, Snodgrassella*, *Neisseria*, *Simonsiella*, and *Uruburuella* (Supplementary Table 5).

A phylogenetic tree constructed using 47 sequences representing highly prevalent *Betaproteobacteria* OTUs from the Belheouane G15 AIL dataset and the "Beta12" clone sequences shows that the clones are closely related to the representative OTUs (Figure 1). A distance matrix constructed using Tamura-Nei genetic distance model (Tamura and Nei, 1993) reveals that seven clone sequences share more than 99% identity with representative

OTU_00001, with the remaining six clones sharing more than 98% identity (Supplementary Table 6). Since the RDP SeqMatch taxonomy was updated in 2020, we conducted a new query to see if the taxonomy of OTU_00001 had changed. We find that with the taxonomy database update, OTU_00001 best matches with unclassified *Neisseriaceae* (S_ab score: 0.88) (Supplement Table 7). We then queried the OTU_00001 sequence using NCBI's MegaBLAST tool[17]. We find OTU_00001 most closely aligns with *Microvirgula aerodenitrificans* (94.5% similar identity). OTU_00001 sequence also aligns with *Snodgrassella alvi*, but with a lower similar identity score of 92.7% (Supplementary Table 8). This is noteworthy, as the MegaBLAST query of our "Beta12" clone sequences found that they also align with *Snodgrassella alvi* and returned a high identity score for this hit.

Several of the genera representing high identity scores for the "Beta12" clone sequences have been identified as mammalian commensals, with some representing important human pathogens. For example, *Neisseria* spp. cause gonorrhea and meningitis in humans[19,20]. Both *Kingella* and *Eikenella* have been found to cause endocarditis[21–23]. *Simonsiella* has been identified as a commensal in the human oral cavity[24,25]. As such, it stands to reason that these bacteria might also be commensal taxa in mice. However, *Microvirgula* and *Snodgrassella*, which are also both members of *Neisseriaceae*, may not be mammalian commensals. Each genus contains only one known species. *Microvirgula aerodenitrificans* is an aerobic and heterotrophic denitrifier; its usual habitat is sludge and waste treatments, but it has also been identified in ponds and canals[26–29]. *Snodgrassella alvi* is an obligate aerobe that has been previously identified in the guts of honeybees and bumble bees[30,31].

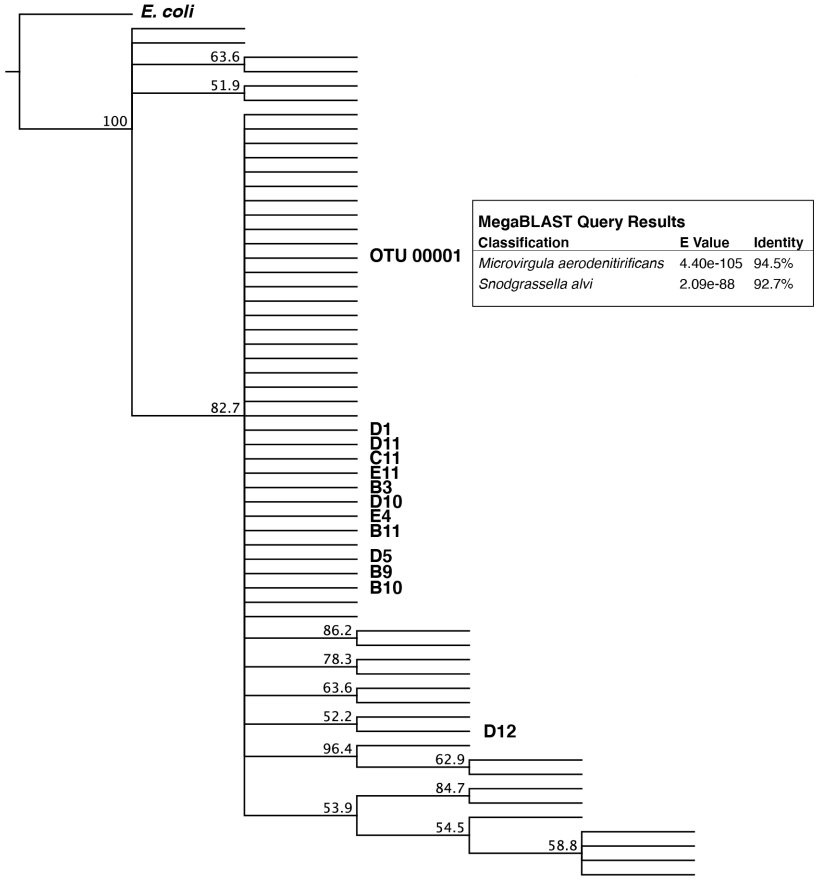


Figure 1. Rooted phylogenetic tree (*E. coli* as outgroup) including 47 sequences representing highly prevalent *Betaproteobacteria* OTUs from the Belheouane G15 AIL dataset and "Beta12" clone sequences inferred using the Neighbor- Joining method[32] in Geneious version 9.1.5 from Biomatters (http://www.geneious.com)[18]. Clone sequences are labelled with a capital letter and number, e.g., "D1."

Genetic distances were computed using the Tamura-Nei genetic distance model[33]. The bootstrap consensus networks inferred from 1000 iterations were taken to represent the genetic distances of the taxa analyzed. The branch labels report consensus support (%). Two MegaBLAST hit results are provided for representative OTU_00001. The e-value is the number of hits of similar quality that can be expected by chance. The smaller the e-value, the better the match. The identity represents the precent similarity between the queried sequence and the hit.

Quantitative gene expression—G15 AIL samples

To further evaluate the association between genotype at the *Nell2*-linked SNP marker and the unclassified *Betaproteobacteria* taxon, we performed quantitative real-time PCR of *Nell2* expression on a subset of G15 AIL samples (n=20 randomly selected among each genotype class) to assess whether gene expression might differ according to genotype (see Methods). We find a non-significant increase in gene expression for the CC genotype at *Nell2* SNP UNC26160173 (chromosome 15) compared to the CT and TT genotypes (Kruskal-Wallis chi-squared = 1.06, $p = 0.59$; Wilcoxon $p > 0.05$; Figure 2A). We then assessed whether the relative abundance of unclassified *Betaproteobacteria* might vary according to *Nell2* expression in this same subset of samples using 16S rRNA sequencing data obtained from Belheouane et *al.*[2]. Interestingly, *Nell2* expression inversely correlates with unclassified *Betaproteobacteria* relative abundance (Spearman's rho = -0.59, $p = 0.036$; Figure 2C) in the G15 AIL samples. We find CC genotype samples contain the lowest relative abundance, heterozygotes contain a moderate amount, and the highest levels are observed within the TT genotype samples (Figure 2B). The differences in relative abundances between groups are not statistically significant (Kruskal-Wallis chi-squared = 1.05, $p = 0.59$; Wilcoxon $p > 0.05$).

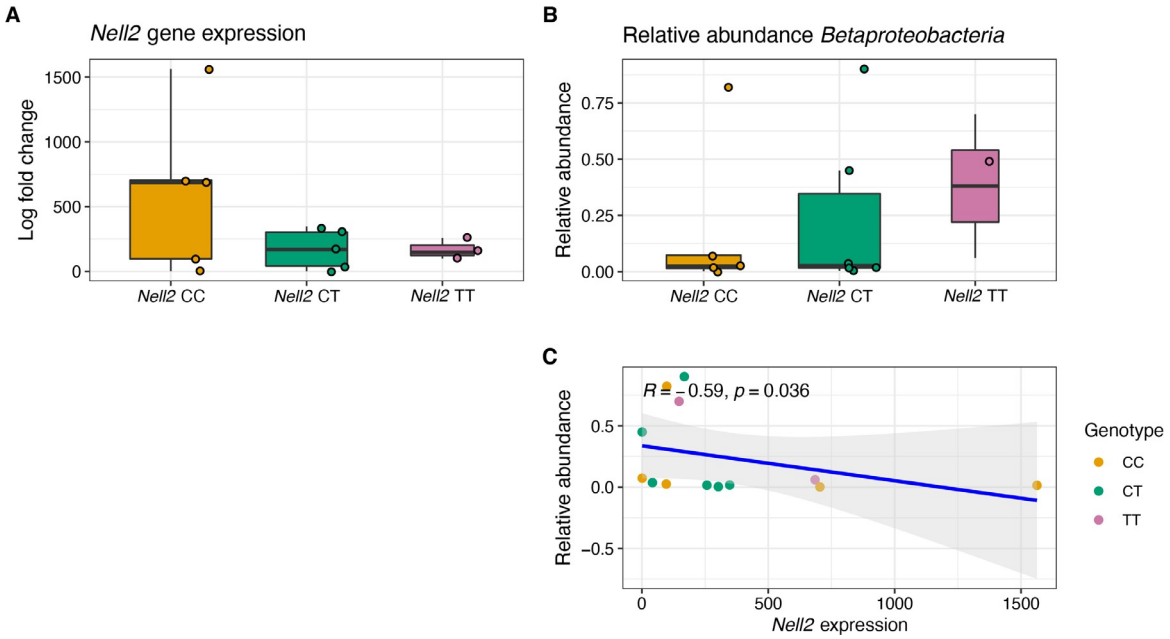


Figure 2. Plots of *Nell2* gene expression and relative abundances of unclassified *Betaproteobacteria* in G15 AIL samples. **A)** Boxplots of *Nell2* gene expression according to examined SNP genotype. **B)** Boxplots of the relative abundances of unclassified *Betaproteobacteria* according to *Nell2* genotype. **C)** Scatterplot of *Nell2* gene expression and unclassified *Betaproteobacteria* relative abundances. For boxplots: Line indicates the median concentration; box shows the interquartile range (IQR), and the whiskers are 1.5x IQR. For scatterplot: blue line shows the regression line with the confidence interval at 95% represented in light grey. R = Spearman's rho. Gold represents *Nell2* CC genotype, green represents *Nell2* CT genotype, and pink represents *Nell2* TT genotype. Reported *p* values are corrected for multiple testing using the Benjamini-Hochberg method[34], when necessary. Summary statistics are provided in Supplementary Table 9.

Skin microbiota composition—*Nell2* mouse samples

Forty-two ear samples were included for DNA/RNA extraction from a mutant mouse line representing *Nell2* wildtype, *Nell2* heterozygous, and *Nell2* knock-out genotypes (see Methods; Supplementary Table 10). We performed 16S rRNA gene amplicon (V1-V2 hypervariable regions) sequencing on the Illumina MiSeq platform using RNA reverse transcribed into cDNA as template. The final analysis included 41 samples consisting of 14 samples for the wildtype and knock-out groups, each containing seven males and seven females, and 13 samples for the heterozygous mouse group, consisting of seven males and six females; one sample was removed from the dataset after quality filtering and processing. In total, we analyzed more than 100,000 sequences, with a normalized sequencing depth of 2,500 sequences per sample. Supplementary table 11 provides the read counts for all quality filtering and processing steps in the analysis.

We first analyzed skin microbial community composition at the phylum- and genus-level (Figure 3; Supplementary Table 12). Overall, we find *Firmicutes*, *Proteobacteria*, *Actinobacteria*, and *Bacteroidetes* represent the most abundant phyla. *Ochrobactrum, Staphylococcus*, *Atopostipes*, *Jeotgalicoccus*, and *Turicibacter* are amongst the most abundant genera. Wilcoxon rank sum analyses reveal that the abundances of phyla and genera do not significantly differ between genotypes (Supplementary Table 12.1). We then assessed skin microbial community composition according to both genotype and sex. Again, we find that the major phyla and genera do not significantly vary between sexes within each genotype (Wilcoxon tests, $p > 0.05$; Figure 3; Supplementary Table 12.2).

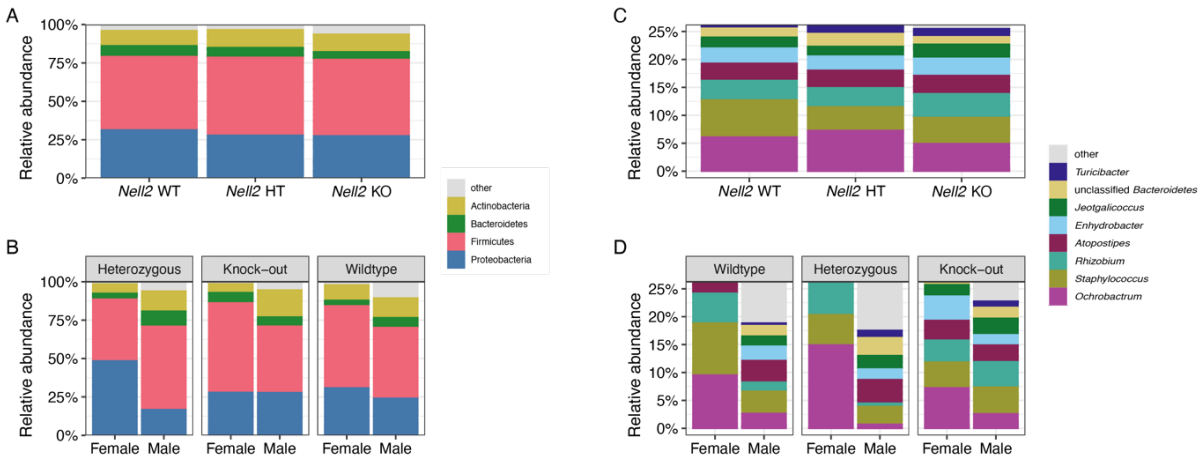


Figure 3. Relative abundances of phyla and genera. **A, C)** The most abundant phyla and genera are shown for each *Nell2* genotype. **B, D)** The most abundant phyla and genera are shown for each *Nell2* genotype and stratified by sex.

Skin microbiota diversity indices—*Nell2* mouse samples

Next, we assessed alpha and beta diversity patterns according to *Nell2* genotype. We first assessed alpha diversity at the amplicon-sequence-variant (ASV) level to investigate the potential relationship between *Nell2* and skin microbiota. Shannon diversity measures both the richness (number of different species) and evenness (how the species are distributed relative to one another) of the bacterial community, whereas the Chao1 index reflects expected species richness. We find no significant differences in Shannon diversity or Chao1 richness between genotypes (Figure 4; Supplementary Table 13). However, after stratifying by sex, we find a significant difference between the sexes in Shannon diversity for both the heterozygous and knock-out genotypes and in Chao1 richness for the heterozygous genotype ($p = 0.039$; $p = 0.039$; $p = 0.024$,

respectively).

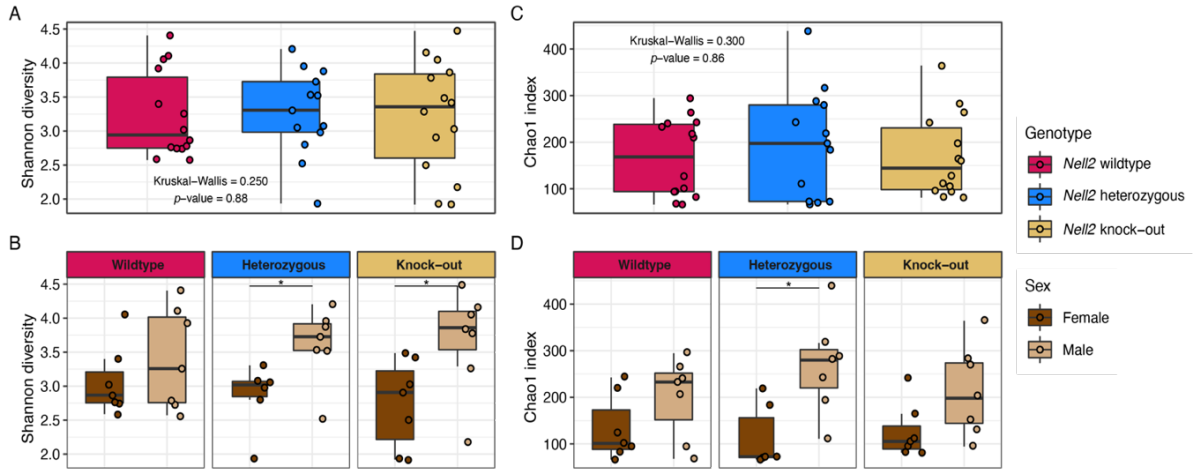


Figure 4. Boxplots of alpha diversity measures. **A, B)** Shannon diversity and Chao1 index between *Nell2* genotypes. **C, D)** Shannon diversity and Chao1 index faceted by sex within *Nell2* genotypes. Summary statistics for boxplots are provided in Supplementary Table 14. For plots B and D: Wilcoxon test (see Methods); *p*-values: * <0.05; ** < 0.01; *** < 0.001. *p*- values were adjusted for multiple testing following the Benjamini and Hochberg method[34], when necessary. Line indicates the median concentration; box shows the interquartile range (IQR), and the whiskers are 1.5x IQR.

To determine whether these findings were affected by how the mice were housed, we calculated a linear mixed model using "maternal ID" as a random term to estimate variability in alpha diversity measures, while controlling for sex and genotype (see Methods). Of note, mice were caged according to maternal ID and thus maternal ID is a proxy for "cage." We find a moderate, significant, sex effect for both Shannon diversity and Chao1 richness, and we reveal that sex and maternal ID are correlated (ANOVA; Shannon: Chi-squared = 11.8, $p = 0.0006$; Chao1: Chi-squared = 12.21, $p = 0.0005$). Notably, both models slightly improve in terms of fit and significance when genotype is removed from the mixed model, further suggesting sex and maternal ID are highly confounded (ANOVA; Shannon: Chi-squared = 12.30, $p = 0.0005$; Chao1 ANOVA Chi-squared = 12.81, $p = 0.0003$). Tables 2 and 3 show the effect sizes for the fixed and random effects of the mixed models as well as the variance explained by the fixed effects (marginal R2) and by the entire model, including random effects (conditional R2), for both Shannon diversity and Chao1, respectively. We find that sex explains about 25% of the variance in Shannon diversity and about 40% of the variance in Chao1 richness when genotype is excluded from the models. The incorporation of genotype into the mixed models does not further explain variability in Shannon diversity. For Chao1, the variance explained increases from 40% to 42% when genotype is included in the model.

Table 2. Shannon Diversity mixed model results						
	Mixed model including genotype			Mixed model excluding genotype		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	3.23	3.02 – 3.43	**<0.001**	3.22	3.03 – 3.42	**<0.001**
Knock-out genotype	0.01	-0.27 – 0.30	0.915			
Wildtype genotype	-0.02	-0.29 – 0.26	0.903			
Male sex	-0.34	-0.53 – -0.14	**0.002**	-0.33	-0.53 – -0.14	**0.001**
Random Effects						
σ^2	0.39			0.37		
τ_{00} Maternal ID	0			0		
ICC	0.01			0.01		
N Maternal ID	8			8		
Observations	41			41		
Marginal R^2 / Conditional R^2	0.229/0.235			0.236/0.245		

Table 3. Chao1 index mixed model results						
	Mixed model including genotype			Mixed model excluding genotype		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	174.48	135.11 – 213.84	**<0.001**	173.39	135.42 – 211.36	**<0.001**
Knock-out genotype	19.46	-14.30 – 53.22	0.25			
Wildtype genotype	-8.52	-41.64 – 24.59	0.605			
Male sex	-44.5	-70.36 – -18.64	**0.001**	-45.14	-70.70 – -19.59	**0.001**
Random Effects						
σ^2	5363.85			5320.9		
τ_{00} Maternal ID	1633.17			1480.78		
ICC	0.23			0.22		
N Maternal ID	8			8		
Observations	41			41		
Marginal R^2 / Conditional R^2	0.244/0.421			0.235/0.401		

To assess potential overall community compositional differences between *Nell2* genotypes, we next performed beta diversity analyses. The Bray-Curtis dissimilarity index is calculated using both presence/absence and abundance data whereas Jaccard index uses only presence/ absence and ignores abundance information. Principal Coordinates Analysis (PCoA) plots reveal substantial overlap between the genotypes, suggesting the genotypes exhibit similar microbial communities (Figure 6; Supplementary Tables 14-15). When we conducted canonical analysis of principal coordinates ('*capscale*')[35], derived from the Bray-Curtis dissimilarities, and constrained by sex, we continue to see significant overlap between the genotypes, suggesting community differences do not exist (Figure 6; permutations = 9999). We find that maternal ID accounts for the largest percentage of variation ($p = 0.01$); the variation explained by genotype is not significant (Supplementary Table 16).

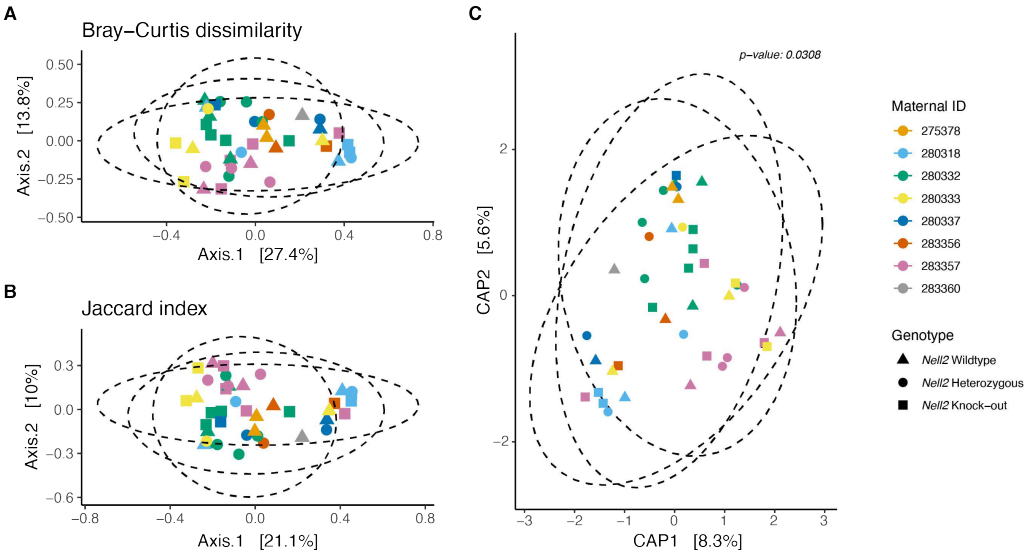


Figure 6. Principle coordinate analyses of skin bacterial beta diversity to visualize differences in microbiota structure according to *Nell2* genotype and sex. **A)** Bray-Curtis dissimilarity and **B)** Jaccard index reveal that the genotypes share similar beta-diversity patterns. **C)** Canonical analysis of principal coordinates ('*capscale*') derived from the Bray-Curtis dissimilarities and constrained by sex reveals significant overlap between the genotypes, suggesting community differences do not exist). We find that maternal ID accounts for the largest percentage of variation (p = 0.01); the variation explained by genotype is not significant. Ellipses represent 95% confidence intervals. Each symbol represents a sample. Triangles = *Nell2* wildtype; circles = *Nell2* heterozygous; squares = *Nell2* knock-out. Bray-Curtis dissimilarity and Jaccard index were calculated in Phyloseq and plotted in ggplot2. The 'vegan' package was used within Phyloseq to conduct the '*capscale*' ordination, for which the '*anova.cc*' function (n=9999 permutations) to assess significance was applied[35,36]. Supplementary Tables 14-16 provide vectors and summary data for the respective beta diversity measures and plots.

Indicator species—*Nell2* mouse samples

To identify potentially important individual taxa within genotypes, we conducted indicator species analyses ('*indicspecies*')[37] at the genus and ASV-levels on a microbiota core defined by a prevalence threshold, whereby a taxon must be present in at one-third of samples for inclusion in the analysis (see Methods). Group comparisons were made according to the presence of at least one copy of the *Nell2* gene (knock-outs versus wildtype and heterozygotes). We find no statistically significant indicator taxa for the knock-out group or the wildtype and heterozygote group after correction for multiple testing. However, at the genus-level, *Enterococcus* is an indicator for the group containing at least one copy of the *Nell2* gene (unadjusted p = 0.032; adj. p = 1.00) and unclassified *Enterobacteriaceae* is an indicator for the knock-out group (unadjusted p = 0.0015; adj. p = 0.19). At the ASV-level, we find thee ASVs with unadjusted p-values less than or equal to a significance threshold of 0.05 for the group containing at least one copy of the *Nell2* gene. These include SV18 (*Enterococcus*; unadjusted p = 0.033, adj. p = 1.00), SV454 (*Clostridium XIVa*; unadjusted p = 0.030; adj. p = 1.00) and SV611 (*Odoribacter*; unadjusted p = 0.026; adj. p = 1.00). For the knock-out group, we find four indicator SVs with unadjusted p-values less than or equal to a significance threshold of 0.05. These include SC317 (*Enterobacteriaceae*; unadjusted p = 0.001; adj. p = 0.163), SV453 (*Lachnospiraceae*; unadjusted p = 0.038; adj. p = 1.00), SV545 (*Actinomyces*; unadjusted p = 0.49; adj. p = 0.038), and SV750 (*Lachnospiraceae*; unadjusted p = 0.022; adj. p = 1.00).

We then took a targeted approach to identify potentially significant differences in taxa previously shown to be associated with *Nell2* or genomic regions containing the *Nell2*. Specifically, we

assessed the prevalence and abundance of unclassified *Betaproteobacteria* and *Herbaspirillum* within the *Nell2* mice samples. *Herbaspirillum* was previously identified in a genome-wide association study using the fourth generation of an AIL susceptible to autoimmune blistering disease to be associated with the genomic region containing *Nell2* on chromosome 15[3]. We find that overall, these taxa are present in relatively few samples; unclassified *Betaproteobacteria* is present in eight samples and *Herbaspirillum* is present in just five samples, and each taxon represents less than 0.5% of the total abundance of skin microbiota. The relative abundances of these taxa are not significantly different between genotypes (Supplementary Table 17).

Next, we assessed the prevalence of *Snodgrassella* and *Microvirgula* in the *Nell2* mice dataset. *Microvirgula* was identified by RDP classifier as a close match to OTU_00001, the representative sequence for unclassified *Betaproteobacteria* in the Belheouane G15 AIL dataset. Both *Microvirgula aerodenitrificans* and *Snodgrassella alvi* aligned with OTU_00001 in a NCBI MegaBLAST query[17]. An NCBI MegaBLAST query of "Beta12" clone sequences aligned with *Snodgrassella alvi* with a high identity score. Therefore, we reasoned that these taxa might significantly differ between genotypes. However, we find that neither genus is present within the *Nell2* mice dataset. Thus, we were unable to verify that *Nell2* genotype associates with skin microbiota.

As we were unable to identify significant indicator taxa, we reannotated the 47 representative OTU sequences for the unclassified *Betaproteobacteria* trait from the G15 AIL dataset. For this, we queried the representative sequences in two taxonomy databases. We find that the Genome Taxonomy Database (GTDB, Release 07-RS207)[38,39] classifies 45 of the 47 sequences within the *Gammaproteobacteria* class, rather than *Betaproteobacteria* (Supplementary Table 18). Sequences annotated beyond the family level were classified as *Parasutterella*, *Turicimonas*, *Mesosutterella*, or *Ralstonia*. We find that the SILVA rRNA taxonomy database (version SILVA SSU 138)[40] annotates 46 of the 47 representative sequences as belonging to class *Gammaproteobacteria* (Supplementary Table 19). Of these, three sequences were annotated to the genus-level and classified as *Gallionella, Sutterella,* and *Parasutterella.* Representative sequence OTU_00001 was classified as belonging to the order *Burkholderiales* by both taxonomy databases and was not annotated further.

Finally, we aligned the 47 representative sequences for the unclassified *Betaproteobacteria* trait from the G15 AIL dataset with sequences derived from the *Nell2* wildtype mouse samples to assess sequence similarities. A phylogenetic tree reveals that 41 of the 47 representative sequences cluster together to create a distinct subclade (Figure 7; Supplementary Table 20). The closest related sequences from the *Nell2* mice dataset include three small subclades made up of ASVs classified as *Neisseria*, *Kingella*, and *Flavobacterium* species.

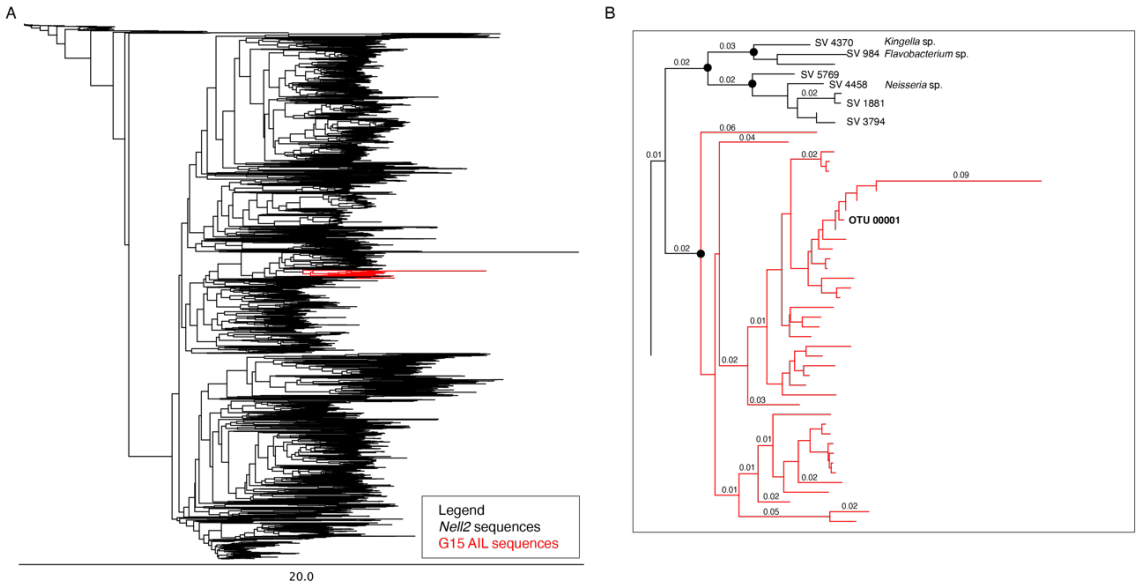


Figure 7. A. Rooted phylogenetic tree (*E. coli* as outgroup) and B. selected subclades of the rooted phylogenetic tree from 47 representative sequences for the unclassified *Betaproteobacteria* trait from the G15 AIL dataset and sequences from the *Nell2* mutant mice dataset inferred using the Neighbor-Joining method (Saitou and Nei, 1987) in Geneious version 9.1.5 from Biomatters (http://www.geneious.com; Kearse et al., 2012). Genetic distances were computed using the Tamura-Nei genetic distance model (Tamura and Nei, 1993). The bootstrap consensus networks inferred from 1000 iterations were taken to represent the genetic distances of the taxa analyzed. Branches are labeled with substitutions per site. Sequences (ASVs) from the *Nell2* mice data are shown in black and sequences (OTUs) from the G15 AIL dataset are colored in red.

Discussion

In this study, we aimed to better characterize the relationship between host genetic variation and the skin microbiota in mice through the analysis of a *Nell2* knock-out strain, as this gene was previously found to associate with unclassified *Betaproteobacteria* in a QTL mapping study by Belheouane and colleagues[2]. We first endeavored to refine the taxonomic classification of the bacterial trait, i.e., unclassified *Betaproteobacteria*, through 16S rRNA gene amplicon sequencing. Then, using a *Nell2* knock-out strain, we analyzed the structure of skin microbiota and searched for bacterial indicators that might significantly differ between *Nell2* wildtype, heterozygous, and knock-out genotypes.

We successfully amplified and cloned a portion of the 16S rRNA gene from Belheouane G15 AIL samples rich in unclassified *Betaproteobacteria*. Clone sequence analysis identified genera from the family *Neisseriaceae* as the most likely bacterial candidates for the unclassified *Betaproteobacteria* trait associating with *Nell2*. Clustering observed in a phylogenetic network comprising representative unclassified *Betaproteobacteria* sequences and our "Beta12" clone sequences further suggest that a genus from *Neisseriaceae* might be driving the *Nell2* gene-microbe association. Interestingly, we find *Nell2* expression significantly negatively correlates with the relative abundance of unclassified *Betaproteobacteria*, suggesting a dose-dependent genotype effect on bacterial variation within the G15 AIL dataset. We do not observe significant differences in unclassified *Betaproteobacteria* abundances between SNP genotypes, however. Notably, *Nell2* expression was low across the sample set, with only 13 of the 60 samples analyzed returning median cycle threshold (Ct) values at or below 45 cycles.

***Nell2* genotype and skin microbiota do not covary**

Within our *Nell2* mutant mice strain, we largely do not observe significant differences in microbial taxa or alpha and beta diversities between the examined genotypes. In brief, the examined genetic variation of *Nell2* does not associate with differences in microbial structure in murine skin. These results are surprising, given findings from previous QTL mapping studies exploring host gene-microbe associations in murine skin. In 2013, Srinivas et al. reported an association between a QTL on chromosome 15 and *Herbaspirillum* in a genome- wide mapping study of the 4th generation of an AIL susceptible to autoimmune blistering disease[3]. The Belheouane 2017 QTL mapping study was a follow-up of this work[2]. Here, the authors used the 15th generation of the same AIL, resulting in highly recombined individuals and over 50,000 informative SNPs, which dramatically reduced the size of the genomic intervals involved. Notably, in some cases, these genomic intervals contained single genes, like the one containing *Nell2* on chromosome 15. Accordingly, we reasoned that the associated taxa, unclassified *Betaproteobacteria* from Belheouane et al. (2017) or *Herbaspirillum* from Srinivas et al. (2013), might differ between the *Nell2* genotypes studied in our knock-out line. However, when we examined the abundances of these candidate taxa, we failed to identify significant differences between the genotypes. Moreover, we were unable to detect the presence of *Snodgrassella* and *Microvirgula*, which were each identified as potential candidates for the unclassified *Betaproteobacteria* trait through the analysis of a 16S rRNA gene clone library.

It is possible that the *Nell2* variation examined in our study is not associated with abundances of bacterial taxa in the skin. However, it is worth noting that the mouse line used in the QTL mapping study was an autoimmune-prone advanced intercross susceptible to skin inflammation. *Nell2* is predominately expressed in neural tissues[7], however it has been found to be differentially expressed in skin within the context of the inflammatory skin disease atopic dermatitis[8]. It is possible that the *Nell2*-microbe association identified in the QTL study is not present in the *Nell2* knock-out due to environmental differences (*e.g.,* different handlers, different mouse facilities). Alternatively, it is possible that there is an indirect or epitasis effect, whereby mutations in other genes must also be present for *Nell2* to affect changes in the skin microbiota. Lastly, it is also possible that the bacterium associated with *Nell2* was not present in the embryo transfer during mouse rederivation.

Changes in bacterial taxonomy and its impact on our study

Our study has other limitations. Since the publication of the QTL mapping study in 2017, taxonomy databases used to classify bacteria have been updated to reflect evolving taxonomy that better represent evolutionary relationships between bacteria and their biochemical properties and the introduction of novel microorganisms. In particular, the class *Betaproteobacteria* has now been reclassified as an order within the *Gammaproteobacteria* class [38,39]. This re-classification explains why OTU_00001 is now classified as belonging to *Gammaproteobacteria* in both the GTDB and SILVA rRNA databases. Interestingly, the updated taxonomic classifications then diverge; GTDB classifies OTU_00001 to the family *Burkholderiaceae*, while the SILVA database classifies OTU_00001 to the family *Neisseriaceae*. In general, *Neisseriaceae* comprises many genera known to be mammalian commensals and important human pathogens, including *Eikenella, Neisseria*, *Kingella*, and *Simonsiella*[21,22,24,25,41–47]. On the other hand, *Burkholderiaceae* contains many genera that have been labelled as common contaminants in microbiome studies[48–50].

Accordingly, when we reannotated the 47 representative sequences for the unclassified *Betaproteobacteria* trait, GTDB classified one of these OTU sequences as *Ralstonia*, which

has been identified as a common contaminant in DNA/RNA extraction kits and microbiome studies[49]. A pairwise alignment between OTU_00001 and the sequence representing *Ralstonia* reveals that the two sequences are 95.5% identical (Supplementary Table 21). Historically, sequences that share >95% identity are typically classified to the same genus[51,52]. The GTDB-toolkit is largely considered a robust classification tool; it provides objective taxonomic assignments based on the phylogeny of genomes sourced from the NCBI Assembly database[39]. Thus, it is possible that the GTDB taxonomic classification for OTU_00001 better represents its true taxonomy. Moreover, unclassified *Betaproteobacteria* and *Herbaspirillum* (family *Oxalobacteraceae*) were present in minute abundances, and we were unable to identify the other bacterial candidates from *Neisseriaceae*, namely, *Microvirgula* and *Snodgrassella*, in our *Nell2* dataset. These findings, in conjunction with updated taxonomic information, suggest that the bacterial trait associated with *Nell2* might rather belong to *Burkholderiaceae*, and not *Neisseriaceae*. In other words, we might have targeted the wrong family for our 16S rRNA gene clone library analysis. Additionally, these data collectively suggest that it is possible that the unclassified *Betaproteobacteria* trait was a spurious finding, perhaps stemming from kit or reagent contamination inherent in microbiome studies[48,53,49,50].

Conclusion

We find that skin microbiota composition does not significantly differ between the examined *Nell2* genotypes. Moreover, we were unable to replicate the finding that the relative abundance of unclassified *Betaproteobacteria* varies according to genotype within a *Nell2* knock-out line. Finally, we reveal evidence suggesting that the unclassified *Betaproteobacteria* trait previously found to be associated with *Nell2* is likely a member of the *Gammaproteobacteria* class and, more specifically, the *Burkholderiaceae* family, and might represent a contaminant. This hypothesis could be explored using ddPCR with primers targeting specific taxa, e.g., *Ralstonia*, to precisely measure bacterial abundances. Future studies aimed at replicating host gene-microbe associations in a murine model are warranted to assess the reliability of previous findings that bacterial taxa in the skin might be significantly influenced by host genetic variation.

Methods

G15 AIL samples from Belheouane et al. (2017)

G15 AIL animal husbandry and sample collection

As previously described by Belheouane et al.[2] in their work identifying gene-microbe interactions in mouse skin microbiota via QTL mapping, MRL/MpJ, NZM2410/J, BXD2/TyJ, and CAST/EiJ mice (Jackson Lab, Maine, USA) were intercrossed in equal strain and sex distributions to create a heterogenous intercross. In brief, all 270 animals were separated according to family, held under pathogen-free conditions at a 12-hour light/dark cycle, and provided with food and water ad libitum. Mice from the 15th generation (G15) of this advanced intercross line (AIL) were sampled at a mean age of 5.9 months. These animal experiments were approved by the "Ministerium für Energiewende, Landwirtschaft, Umwelt und ländliche Räume des Landes Schleswig-Holstein" in Kiel, Germany (reference number: V 312–72241. 122–5 (12-2/09)).

The dissection process and sampling steps were previously performed by Belheouane et al.[2]. In summary, the authors obtained an identical region from the left ear of each mouse. The samples were snap frozen and stored at -80°C until further processing. Dissection tools were carefully sterilized by flaming 70% ethanol and working surfaces and pipettes were cleaned with RNase AWAY® (Thermo Fischer Scientific). Total DNA and RNA were extracted simultaneously using the AllPrep DNA/RNA Qiagen kit, with an additional 2-hour room temperature incubation step after homogenization to increase the nucleic acid dissolution in the RLT buffer. Extracted RNA was treated with DNase (RNase-Free DNase Qiagen, stock solution concentration) and cDNA synthesis was performed using High-Capacity cDNA Reverse Transcription Kits (Applied Biosystems). The purity of RNA was verified by negative reverse transcriptase PCR (without

transcriptase) and by agarose gel electrophoresis. Samples and data obtained from Belheouane et al.[2] for the current study are referred to as "G15 AIL samples" or "G15 AIL dataset."

Bacterial candidate identification—G15 AIL samples

Primer design was conducted using the NCBI primer-BLAST tool Primer3[13]. Primer characteristics were analyzed using ThermoScientific™ Multiple Primer Analyzer tool[62]. The literature search was conducted in the NCBI database for 16S rRNA primers targeting the *Betaproteobacteria* class. Bacterial reference strains were obtained from The Leibniz Institute DSMZ (German Collection of Microorganisms and Cell Cultures). The reference strains *Microvirgula aerodenitrificans* (DSMZ 15089), *Neisseria gonorrhoeae* (DSMZ 9188), *Nitrosomonas europaea* (DSMZ 28437), and *Sutterella wadsworthensis* (DSMZ 14016) were selected to assess the performance of primers across a range of taxa within the *Betaproteobacteria* class. Additionally, *E. coli* (DSMZ 30083) was selected as a reference strain to assess the specificity of primer sets.

PCRs were conducted in a 30-µl volume containing 100ng cDNA using Phusion Hot Start II DNA High-Fidelity DNA Polymerase (Finnzymes, Espoo, Finland). The cycling conditions were as follows: initial denaturation for 2 minutes at 95°C; 35 cycles of 30 seconds at 95°C, 30 seconds at 55°C and 3 minutes at 73°C; final extension for 10 minutes at 72°C. PCR products were loaded on a 1.5% agarose gel to confirm 16S rRNA gene amplicon bands. Products with the expected band size were selected and cloned using the CloneJet PCR kit from ThermoScientific and One Shot TOP10 Chemically Competent *E. coli* from Invitrogen, followed by a purification step with MiniPrep (Qiagen). The presence of DNA fragment insertion was confirmed by direct PCR amplification with pJet1.1 forward and reverse primers. The inserted DNA fragment was then sequenced using BigDye® Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems). For quality assurance, the sequences were manually edited and aligned using Geneious version 9.1.5 from Biomatters (http://www.geneious.com)[18]. Chimeric sequences were identified and removed from the analysis using the Find Chimeras web tool from DECIPHER (http://DECIPHER.cee.wisc.edu)[63]. Assembled consensus sequences were then aligned and compared to representative OTU sequences for the unclassified *Betaproteobacteria* trait. The sequences were queried within Geneious Prime using NCBI's basic local alignment search tool (MegaBLAST)[17] (http://blast.ncbi.nlm.nih.gov/Blast.cgi), which accesses the GenBank® database (http://www.ncbi.nlm.nih.gov/genbank/)[64]. Taxonomy of trimmed clone sequences was additionally defined in RDP SeqMatch (v.3, release 11_6)[12]. Results reported are based on the highest SeqMatch score (S_ab).

Phylogenetic trees of the sequenced clones and reference strains were inferred using the Neighbor-Joining method[32] in Geneious version 9.1.5 from Biomatters (http://www.geneious.com)[18]. Genetic distances were computed using the Tamura-Nei genetic distance model[33]. The bootstrap consensus trees inferred from 1000 iterations were taken to represent the genetic distances of the analyzed taxa.

Quantitative gene expression—G15 AIL samples

Reverse transcription quantitative PCR was carried out for quantifying *Nell2* gene expression on a subset of 60 selected RNA samples from the Belheouane G15 AIL dataset, with 20 samples randomly selected for each genotype (Supplementary Table 2)[2]. cDNA synthesis was performed using High-Capacity cDNA Reverse Transcription Kits (Applied Biosystems). RNA purity was checked by negative reverse transcriptase PCR (without transcriptase) and by agarose gel electrophoresis. Real-time quantification PCR of *Nell2* expression was performed using *Nell2* mouse qPCR primer pair NM_016743 ('5- GAACCACCTACCGAGAGTCTGA-3' and 5'-CTCCTTACAGCACTTGCCATCC-3'; OriGene Technologies, Germany). Real-time qPCR was conducted in a volume of 10-µl on a PikoReal2 Real-Time PCR system using 96-well plates with

three technical replicates for each sample. Each PCR mixture consisted of 5.0-µl PowerUp SYBR PCR Master Mix (Applied Biosystems), 0.25-µl of each primer (10 uM), 3.5-µl of water, and 1-uL of cDNA template. The PCR program was as follows: (i) initial step at 95° C for 10 minutes, (ii) 50 cycles of denaturation at 95°C for 15 sec and annealing/extension at 60°C for 1 minute, and 1 cycle at 60°C for 30 seconds and a melt ramp from 60°C to 95°C. The relative transcript levels of *Nell2* were calculated for each sample using the 2−ΔΔCT method relative to a reference gene (*hprt1*). The final analysis of real-time quantification PCR of *Nell2* expression was conducted for 13 of the 60 samples investigated, after selecting data for samples with three successful replicates for both the *Nell2* gene and housekeeping gene *hprt1* and for which median cycle threshold (Ct) values for *Nell2* were at or below 45 cycles. As the data were not normally distributed, unpaired Wilcoxon rank sum tests performed to compare genotype groups with significance reported for $p < 0.05$.

***Nell2* mutant mice**

***Nell2* mutant mice husbandry and sample collection**

Cryopreserved *Nell2* mutant embryos were purchased from RIKEN BRC in Japan and maintained on C57BL/6J background in the animal facility at the University of Lübeck, Germany under permit number 2018-12-20-Ibrahim for animal sacrifice and organ collection. Offspring were caged according to gender and maternal ID. All mice were kept in ventilated cages with 12 h light-dark cycle and with ad libitum water and a standard chow diet. The approval for mouse husbandry and all animal experiments were approved by the University of Lübeck, Germany. Supplementary Table 10 provides an overview of mice and respective breeding and sample information. Samples and data derived from this mouse line are specified as "*Nell2* mice samples" or "*Nell2* mice dataset."

Genotyping was carried out using genomic DNA extracted from ear samples; 100-µl of 50 mM NaOH were added to the ear punch biopsies and incubated for 1 h at 95°C while shaking. The samples were then vortexed, 10-µl; 1M Tris-HCl added for neutralization, and then vortexed again. The primer pair for the mutant mouse strain included Nell2-OR238-R1 (5'-AATGTGGTTGTTCTACAAGAGCAGAAAAGG -3') and T/BAL (5'-CTTGTGTCATGCACAAAGTAGATGTCC-3') and the pair Nell2-OR238-F1 (5'-TCCCAGATGTCACATAGGAGCAGGAAGTAC-3') and Nell2-OR238-R1 (5'-AATGTGGTTGTTCTACAAGAGCAGAAAAGG -3') for the wild mouse strain at 10 pmol/ µl concentrations. PCR was conducted in 25-µl volume using DreamTaq PCR Master Mix (ThermoFischer Scientific) under the following conditions: initial denaturation for 3 min at 94°C; 35 cycles of 30 s at 94°C, 30 s at 55°C and 60 s at 72°C; final extension for 5 min at 72°C. PCR products were loaded on a 1.0% agarose gel in TAE buffer to confirm the gene bands (383 bp long for wild type and 244 bp for the mutant strain). Animals were sacrificed between two and four months of age via CO2 asphyxiation, followed by cervical dislocation. Mice were weighed prior to dissection. The following tissues were sampled:

Blood: Cardiac blood was sampled and stored in EDTA tubes. The tubes were spun down, plasma transferred into fresh 1.5 ml tubes and stored at -20°C until further processing. The blood cells were discarded.
Brain: The brain was removed and divided along the corpus callosum. Half of the brain was snap frozen by immersion in liquid nitrogen and stored at -80°C until further processing. The other half of the brain was placed into 4% PFA until further processing for histochemistry.
Spleen and mesenteric lymph nodes: The spleen was divided into two sections; one section was preserved in 4% PFA and the other in RNA*later*. A sample from the mesenteric lymph nodes was divided into two sections; one section was preserved in 4% PFA and the other in RNA*later*.
Colon/ Cecum: The entire colon was removed; the colon was transversally divided, with one half of the colon was stored in RNA*later* and the other half in 4% PFA. Cecal content was collected

on ice and transferred to -80°C until further processing. The cecum was cut vertically in half, with one half of the cecum stored in RNA*later* and the other half in 4% PFA.
Ear: Both ears were removed from the mice. The right ear was divided into three pieces. Identical regions of the right ear from each mouse were stored in RNA*later*, a medium of 20% glycerol-BHI, and in 4% PFA, respectively. The left ear was preserved in 4% PFA.
Dorsum: The fur was removed from the dorsum of each mouse using an electric shaver that was sterilized between mice. Then, the dorsal skin was separated from the connective tissue and divided into four pieces, with one piece stored in RNA*later*, one in a medium of 20% glycerol-BHI, one in a medium of 20% glycerol PBS, and the last in 4% PFA until further processing.

Samples preserved in RNA*later* were left overnight at 4°C, spun down, and the supernatant depleted before being stored at -80°C until further processing. All samples were stored at -80°C until further processing. To avoid cross-contamination, all dissection tools were rinsed with water and then sterilized with 70% ethanol and RNase AWAY between mice.

DNA/RNA extraction, 16S rRNA gene sequencing—*Nell2* mice samples
DNA and RNA were extracted simultaneously from 42 mouse ears that were stored in 20% glycerol-BHI at -80°C. Extractions were carried out using the Qiagen AllPrep DNA/RNA kit according to manufacturer's instructions, using β-mercaptoethanol as the reducing agent. The working surface and tools were treated with RNase AWAY® (Thermo Fischer Scientific) to prevent cross-contamination. An additional two-hour incubation step was included after homogenization to increase the dissolution of the nucleic acids in the RTL buffer, as described by Belheouane et al.[2] RNA was treated with DNase (RNase-Free DNase Qiagen, stock solution concentration) as previously described[2]. cDNA synthesis was performed using High-Capacity cDNA Reverse Transcription Kits (Applied Biosystems). RNA purity was checked by a negative reverse transcriptase PCR control and by agarose gel electrophoresis.

The V1-V2 region of the 16S rRNA gene was amplified using RNA reverse transcribed into cDNA as template and the 27F-338R primer pair (5'-**CTATGCGCCTTGCCAGCCCGCTCAG***TC*AGAGTTTGATCCTGGCTCAG-3' and 5'-**CGTATCGCCTCCCTCGCGCCATCAG**XXXXXXXXXX*CA*TGCTGCCTCCCGTAGG AGT-3'; bold = adapter sequence, italics = two-base linker sequence, XXX = ten-base multiplex identifier, underline = 27F–338R). PCR was conducted in a 25-µl volume containing 100ng DNA using Phusion Hot Start II DNA High-Fidelity DNA Polymerase (Finnzymes, Espoo, Finland). The cycling conditions were as follows: initial denaturation for 30 s at 98°C; 30 cycles of 9 s at 98°C, 60 s at 50°C and 90 s at 72°C; final extension for 10 min at 72°C. PCR products were loaded on a 1.5% agarose gel to confirm and quantify the 16S rRNA gene bands, extracted with the GeneJET Gel Extraction Kit (Thermo Scientific, US) and quantified with the Quant-iT dsDNA HS Assay Kit on a Qubit fluorometer (Invitrogen, Darmstadt, Germany). PCR product concentrations were first quantified on an agarose gel using image analysis software (Bio-Rad). After quantification, products were combined accordingly to make equimolar sub-pools. The sub-pools were then extracted from agarose gel using the Qiagen MinElute Gel Extraction Kit and quantified with the Quant- iTTM dsDNA BR Assay Kit on a Qubit fluorometer (Invitrogen). Finally, sub-pools were combined into one equimolar pool for each library. Pools were further purified using AMPure® Beads (Agencourt), and libraries were run on an Agilent Bioanalyzer prior to sequencing, as recommended by Illumina. The libraries were sequenced on a MiSeq using the MiSeq Reagent Kit v3 600 cycles sequencing chemistry (Illumina, CA, US).

Data processing and analyses—*Nell2* mutant mice samples
Data processing and statistical analyses were conducted in R (version 4.0.5). 16S rRNA gene amplicon sequences were processed using DADA2 (version 1.16.0), resulting in ASV abundance tables[65]. "Decontam" (version 1.8.0) was used within Phyloseq (version 1.32.0) to identify

potential contaminant ASVs first according to frequency method, with a strict parameter threshold of 0.1, and then followed by the prevalence method with the strict threshold parameter of 0.5; 315 ASV sequences in total were identified as contaminants and subsequently discarded[36,60,61]. Following recommendations of Weyrich *et al.*[66], ASVs belonging to families *Halomonadaceae* and *Shewanellaceae* were removed. To normalize sequencing coverage, we calculated rarefaction curves to determine sampling threshold; random sub-sampling to 2,500 sequences per sample was performed. Four samples did not meet the 2,500 sequences coverage threshold and were excluded from the analysis.
Additionally, 5673 ASVs were removed because they were no longer present in any sample after random sub-sampling. Taxonomic assignment of ASVs was completed in DADA2 with the Bayesian classifier using the NR Silva database training set, version 138[40].

Alpha diversity measures were calculated for Shannon diversity and Chao1 richness using vegan (version 2.5-6)[35]. As data were not normally distributed, group comparisons were conducted using unpaired Wilcoxon rank sum tests. Beta diversity measures were assessed using the Bray-Curtis dissimilarity index and the Jaccard index in Phyloseq (version 1.32.0)[36]. The vegan package (version 2.5-6) was additionally used to conduct a constrained analysis of principal coordinates ('*capscale'*), a hypothesis-driven ordination that limits the separation of the communities based on the variable tested, for which the 'anova.cca' function was applied to assess significance[35]. Linear mixed models were conducted in R using the function "lmer" from the lme4 package (version 0.9975-3)[67]. "Maternal ID" was defined as a random term and "genotype" and "sex" were defined as fixed terms. Between group relative abundances at the phylum and genus levels were calculated in Phyloseq (version 1.32.0) and compared using unpaired Wilcoxon rank sum tests[36]. Corrections for multiple testing were performed according to Benjamini and Hochberg method[34].

Indicator species analysis was applied using indicspecies (version 1.7.9) with the "r.g." function[68] and 99,999 permutations on a microbial core defined by ASVs or genera classified to the genus-level that are present in at least 33% of all samples. Group comparisons were conducted according to the presence of at least one copy of the *Nell2* gene, i.e., knock-out mice compared to both wildtype and heterozygous together. Significant indicator genera and ASVs were selected after correction of *p*-values for multiple testing using the Benjamini and Hochberg method[34].

Multiple sequence alignment was performed using Geneious version 9.1.5 from Biomatters (http://www.geneious.com)[18]. Phylogenetic networks were inferred using the Neighbor- Joining method[32] in Geneious version 9.1.5 from Biomatters (http://www.geneious.com)[18]. Genetic distances were computed using the Tamura-Nei genetic distance model[33]. The bootstrap consensus networks inferred from 1000 iterations were taken to represent the genetic distances of the taxa analyzed. Sequences were queried within Geneious Prime using NCBI's basic local alignment search tool (MegaBLAST)[17] (http://blast.ncbi.nlm.nih.gov/Blast.cgi), which accesses the GenBank® database (http://www.ncbi.nlm.nih.gov/genbank/)[64]. Taxonomy of sequences was additionally defined in RDP SeqMatch (v.3, release 11_6)[12]. Results reported are based on the highest SeqMatch score (S_ab).

Funding sources: This study was carried out as part of the Research Training Group 'Genes, Environment and Inflammation,' with financial support by the Deutsche Forschungsgemeinschaft (GRK1743, GRK1743/2).

Data availability
The data for this study are available upon request.

Acknowledgements
We would like to thank Katja Cloppenborg-Schmdit and Yasmin Claussen for excellent technical assistance. Thank you to Misa Hirose and Adina-Malin Tietje for assistance with *Nell2* mice breeding and sampling. Thank you to Dr. Malte C. Rühlemann and Dr. Shauni Doms for statistical and analytical expertise.

Competing interests
The authors state no conflict of interest.

Author contributions
Conceptualization: JFB, SI; **Data curation:** BMH (*Nell2* data), MB (G15 AIL data), MH (*Nell2* sampling), AMT (*Nell2* sampling, *Nell2* genotyping); **Study (experiment) conduct:** BMH, MH, AMT; **Formal analysis:** BMH; **Funding acquisition:** JFB, SI; **Investigation:** BMH, AMT; **Methodology:** BMH, JFB, MH, SI; **Project administration:** JFB, SI, MH; **Resources:** BMH, JFB, MH, SI; **Software:** BMH; **Supervision:** MH, JFB, SI; **Validation:** BMH, JFB; **Visualization:** BMH; **Writing–Original draft preparation:** BMH; **Writing– Review and editing:** all authors

References

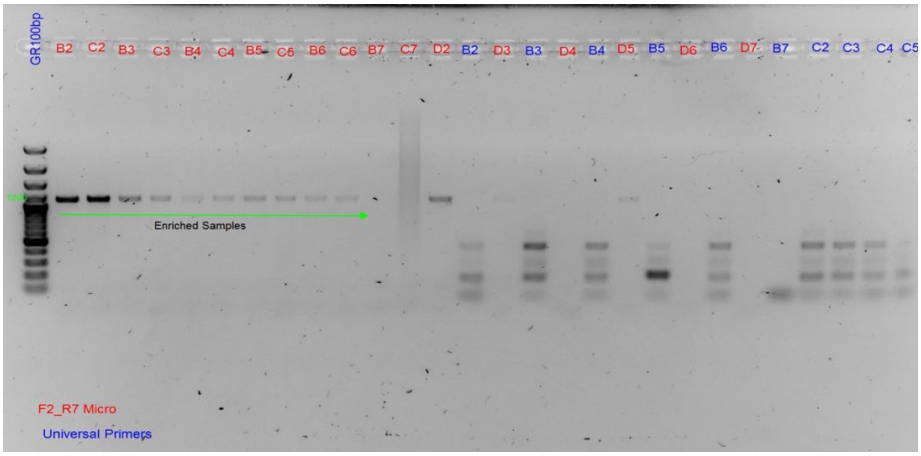
1. Grice, E. A. & Segre, J. A. The skin microbiome. *Nat Rev Microbiol* 9, 244–253 (2011).
2. Belheouane, M., Gupta, Y., Künzel, S., Ibrahim, S. & Baines, J. F. Improved detection of gene-microbe interactions in the mouse skin microbiota using high-resolution QTL mapping of 16S rRNA transcripts. *Microbiome* 5, 59 (2017).
3. Srinivas, G. *et al.* Genome-wide mapping of gene–microbiota interactions in susceptibility to autoimmune skin blistering. *Nat Commun* 4, 1–7 (2013).
4. Kong, H. H. Skin microbiome: genomics-based insights into the diversity and role of skin microbes. *Trends Mol Med* 17, 320–328 (2011).
5. Ley, R. E., Peterson, D. A. & Gordon, J. I. Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* 124, 837–848 (2006).
6. Turnbaugh, P. J. *et al.* The human microbiome project: exploring the microbial part of ourselves in a changing world. *Nature* 449, 804–810 (2007).
7. Kim, D. H. *et al.* Neural Epidermal Growth Factor-Like Like Protein 2 (NELL2) Promotes Aggregation of Embryonic Carcinoma P19 Cells by Inducing N-Cadherin Expression. *PLoS One* 9, e85898 (2014).
8. Kamsteeg, M. *et al.* Molecular diagnostics of psoriasis, atopic dermatitis, allergic contact dermatitis and irritant contact dermatitis. *Br J Dermatol* 162, (2010).
9. Berger, A. Th1 and Th2 responses: what are they? *BMJ* 321, 424 (2000).
10. Kamsteeg, M. *et al.* Type 2 Helper T-Cell Cytokines Induce Morphologic and Molecular Characteristics of Atopic Dermatitis in Human Skin Equivalent. *AM J Pathol* 178, 2091–2099 (2011).
11. Urashima, R. & Mihara, M. Cutaneous nerves in atopic dermatitis. *Virchows Archiv* 432, 363–370 (1998).
12. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl Environ Microbiol* 73, 5261–5267 (2007).
13. Ye, J. *et al.* Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* 13, 134 (2012).
14. Lane, D. J. 16S/23S rRNA sequencing. In *Nucleic acid techniques in bacterial systematics* (eds. Stackebrant, E. & Goodfellow, M.) 115–175 (John Wiley & Sons, Inc, 1991).
15. Mühling, M., Woolven-Allern, J., Murrell, J. C. & Joint, I. Improved group-specific PCR primers for denaturing gradient gel electrophoresis analysis of the genetic diversity of complex microbial communities. *ISME J* 2, 379–392 (2008).
16. Cole, J. R. *et al.* The Ribosomal Database Project (RDP-II): sequences and tools for high-throughput rRNA analysis. *Nucleic Acids Res* 33, D294-296 (2005).
17. Morgulis, A. *et al.* Database indexing for production MegaBLAST searches. *Bioinformatics* 24, 1757–1764 (2008).
18. Kearse, M. *et al.* Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649 (2012).
19. McSheffrey, G. G. & Gray-Owen, S. D. Chapter 82 - Neisseria gonorrhoeae. in *Molecular Medical Microbiology (Second Edition)* (eds. Tang, Y.-W., Sussman, M., Liu, D., Poxton, I. & Schwartzman, J.) 1471– 1485 (Academic Press, 2015). doi:10.1016/B978-0-12-397169-2.00082-2.
20. Rouphael, N. G. & Stephens, D. S. Neisseria meningitidis: Biology, Microbiology, and Epidemiology.
21. *Methods Mol Biol* 799, 1–20 (2012).
22. Nordholm, A. C., Vøgg, R. O. B., Permin, H. & Katzenstein, T. Eikenella corrodens endocarditis and liver abscess in a previously healthy male, a case report. *BMC Infect Dis* 18, 35 (2018).

23. Principi, N. & Esposito, S. Kingella kingae infections in children. *BMC Infect Dis* 15, 260 (2015).
24. Patrick, W. D., Brown, W. D., Ian Bowmer, M. & Sinave, C. P. Infective endocarditis due to Eikenella corrodens: Case report and review of the literature. *Can J Infect Dis* 1, 139–142 (1990).
25. Carandina, G., Bacchelli, M., Virgili, A. & Strumia, R. Simonsiella filaments isolated from erosive lesions of the human oral cavity. *J Clin Microbiol* 19, 931–933 (1984).
26. Caselli, E. *et al.* Defining the oral microbiome by whole-genome sequencing and resistome analysis: the complexity of the healthy picture. *BMC Microbiol* 20, 120 (2020).
27. Patureau, D. *et al.* Microvirgula aerodenitrificans gen. nov., sp. nov., a new gram-negative bacterium exhibiting co-respiration of oxygen and nitrogen oxides up to oxygen-saturated conditions. *Int J Syst Bacteriol* 48, 775–782 (1998).
28. Patureau, D. *et al.* Biological nitrogen removal in a single aerobic reactor by association of a nitrifying ecosystem to an aerobic denitrifier, Microvirgula aerodenitrificans. *Journal of Molecular Catalysis B: Enzymatic* 5, 435–439 (1998).
29. Patureau, Zumstein, Delgenes, & Moletta. Aerobic Denitrifiers Isolated from Diverse Natural and Managed Ecosystems. *Microb Ecol* 39, 145–152 (2000).
30. Patureau, D. *et al.* Combined phosphate and nitrogen removal in a sequencing batch reactor using the aerobic denitrifier, Microvirgula aerodenitrificans. *Water Res* 35, 189–197 (2001).
31. Kwong, W. K. & Moran, N. A. Cultivation and characterization of the gut symbionts of honeybees and bumble bees: description of Snodgrassella alvi gen. nov., sp. nov., a member of the family Neisseriaceae of the Betaproteobacteria, and Gilliamella apicola gen. nov., sp. nov., a member of Orbaceae fam. nov., Orbales ord. nov., a sister taxon to the order 'Enterobacteriales' of the Gammaproteobacteria.
32. Int J Syst Evol Microbiol 63, 2008–2018 (2013).
33. Kwong, W. K., Engel, P., Koch, H. & Moran, N. A. Genomics and host specialization of honeybee and bumble bee gut symbionts. *Proc Natl Acad Sci USA* 111, 11509–11514 (2014).
34. Saitou, N. & Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees.
35. *Mol. Biol. Evol.* 4, 406–425 (1987).
36. Tamura, K. & Nei, M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 10, 512–526 (1993).
37. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 57, 289–300 (1995).
38. Oksanen, J. et al. Community Ecology Package. (2005).
39. McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. in *PloS One* (2013). doi:10.1371/journal.pone.0061217.
40. De Caceres, M. & Legendre, P. Associations between species and groups of sites: indices and statistical inference. *Ecology* (2009).
41. Parks, D. H. *et al.* A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol* 36, 996–1004 (2018).
42. Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P. & Parks, D. H. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* 36, 1925–1927 (2020).
43. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41, D590-596 (2013).
44. Decker, M. D. Eikenella corrodens. *Infect Control* 7, 36–41 (1986).

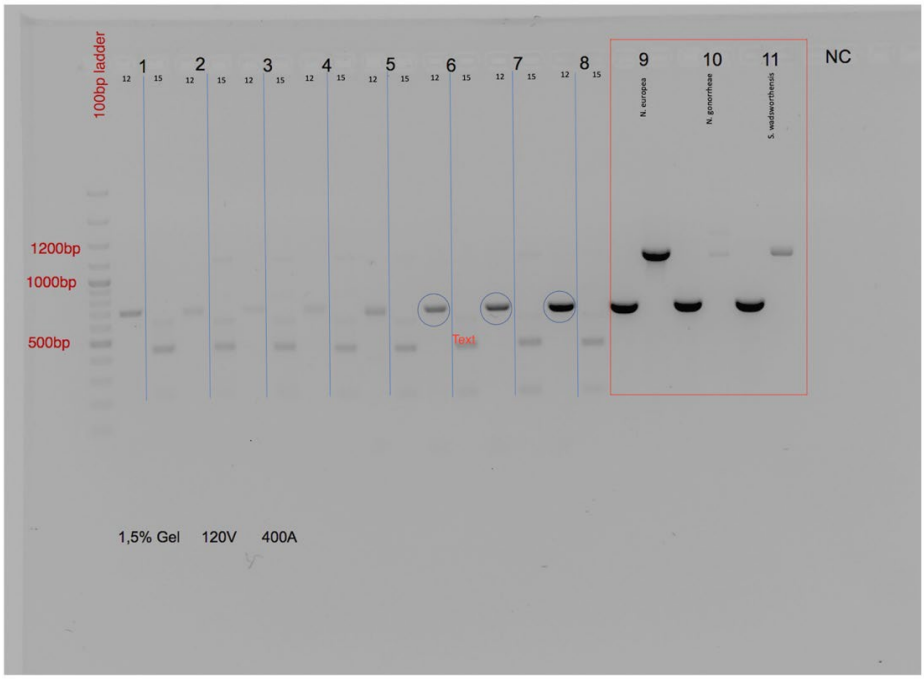
45. Dewhirst, F. E. *et al.* The Human Oral Microbiome. *J Bacteriol* 192, 5002–5017 (2010).
46. Dorey, R. B., Theodosiou, A. A., Read, R. C. & Jones, C. E. The nonpathogenic commensal Neisseria: friends and foes in infectious disease. *Curr Opin Infect Dis* 32, 490–496 (2019).
47. Kim, W. J. *et al.* Commensal Neisseria Kill Neisseria gonorrhoeae through a DNA-Dependent Mechanism.
48. *Cell Host Microbe* 26, 228-239.e8 (2019).
49. Laumen, J. G. E. *et al.* Antimicrobial susceptibility of commensal Neisseria in a general population and men who have sex with men in Belgium. *Sci Rep* 12, 9 (2022).
50. Morse, S. A. Neisseria, Moraxella, Kingella and Eikenella. in *Medical Microbiology* (ed. Baron, S.) (University of Texas Medical Branch at Galveston, 1996).
51. Yagupsky, P., Porsch, E. & St Geme, J. W. Kingella kingae: an emerging pathogen in young children.
52. *Pediatrics* 127, 557–565 (2011).
53. Salter, S. J. *et al.* Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biology* 12, 87 (2014).
54. Eisenhofer, R. *et al.* Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. *Trends Microbiol* 27, 105–117 (2019).
55. Olomu, I. N. *et al.* Elimination of "kitome" and "splashome" contamination results in lack of detection of a unique placental microbiome. *BMC Microbiology* 20, 157 (2020).
56. Johnson, J. S. *et al.* Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun* 10, 5029 (2019).
57. Schloss, P. D. & Handelsman, J. Introducing DOTUR, a Computer Program for Defining Operational Taxonomic Units and Estimating Species Richness. *Appl Environ Microbiol* 71, 1501–1506 (2005).
58. Glassing, A., Dowd, S. E., Galandiuk, S., Davis, B. & Chiodini, R. J. Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. *Gut Pathog* 8, 24 (2016).
59. Sze, M. A., Abbasi, M., Hogg, J. C. & Sin, D. D. A Comparison between Droplet Digital and Quantitative PCR in the Analysis of Bacterial 16S Load in Lung Tissue Samples from Control and COPD GOLD 2. *PLoS One* 9, e110351 (2014).
60. Abellan-Schneyder, I., Schusser, A. J. & Neuhaus, K. ddPCR allows 16S rRNA gene amplicon sequencing of very small DNA amounts from low-biomass samples. *BMC Microbiology* 21, 349 (2021).
61. Quan, P.-L., Sauzade, M. & Brouzes, E. dPCR: A Technology Review. *Sensors (Basel)* 18, 1271 (2018).
62. Vogelstein, B. & Kinzler, K. W. Digital PCR. *Proc Natl Acad Sci U S A* 96, 9236–9241 (1999).
63. Gobert, G. *et al.* Droplet digital PCR improves absolute quantification of viable lactic acid bacteria in faecal samples. *J Microbiol Methods* 148, 64–73 (2018).
64. Maheshwari, Y., Selvaraj, V., Hajeri, S. & Yokomi, R. Application of droplet digital PCR for quantitative detection of Spiroplasma citri in comparison with real time PCR. *PLoS One* 12, e0184751 (2017).
65. Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A. & Callahan, B. J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* 6, 226 (2018).
66. Karstens, L. *et al.* Controlling for Contaminants in Low-Biomass 16S rRNA Gene Sequencing Experiments.
67. *mSystems* 4, e00290-19, /msystems/4/4/mSys.00290-19.atom (2019).
68. Breslauer, K. J., Frank, R., Blöcker, H. & Marky, L. A. Predicting DNA duplex stability from the base sequence. *Proc Natl Acad Sci USA* 83, 3746–3750 (1986).
69. Wright, E. S., Yilmaz, L. S. & Noguera, D. R. DECIPHER, a Search-Based Approach to

Chimera Identification for 16S rRNA Sequences. *Appl Environ Microbiol* 78, 717–725 (2012).
70. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J Mol Biol* 215, 403–410 (1990).
71. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*
72. 13, 581–583 (2016).
73. Weyrich, L. S. *et al.* Laboratory contamination over time during low-biomass sample analysis. *Mol Ecol Resour* 19, 982–996 (2019).
74. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67, 1–48 (2015).
75. Cáceres, M. D. & Legendre, P. Associations between species and groups of sites: indices and statistical inference. *Ecology* 90, 3566–3574 (2009).

Supplementary Figures



Supplementary Figure 1. Gel electrophoresis image showing PCR amplification of a portion of the 16S rRNA gene in cDNA derived from the G15 AIL dataset (Belheouane et al. 2017) using the novel primer pair F2_micro and R7_micro. Samples shown in red represent those amplified using the novel primer pair F2_micro and R7_micro (see Methods). Samples shown in blue represent those amplified using 16S rRNA universal primers 27F and 1392R.[14] Importantly, sample "B7" is type strain *Microvirgula aerodenitrificans* (DSMZ 15089), "C7" is a PCR negative control, and "D7" is *E. coli* type strain (DSMZ 30083). We find that the novel primer pair successfully amplified cDNA samples and the *Microvirgula* type strain. Negative controls were as expected. *E. coli* was not amplified with the novel primer pair, suggesting primer specificity.



Supplementary Figure 2. Gel electrophoresis image showing PCR amplification of a portion of the 16S rRNA gene in cDNA derived from the G15 AIL dataset[2] using the primer pair "Beta12" (see Methods). Samples amplified by Beta12 primers include wells labelled one through eleven. The samples in the remaining rows were amplified by a different primer set (not reported). Samples nine, ten, and eleven include the type strains *Neisseria gonorrhoeae* (DSMZ 9188), *Nitrosomonas europaea* (DSMZ 28437), and *Sutterella wadsworthensis* (DSMZ 14016). NC is the negative control.

Conclusion

The present work characterizes skin microbiota in the context of inflammatory disease. In bullous pemphigoid (BP), an autoimmune blistering disease, we find significant differences in skin microbiota compositions between patients with the disease compared to that of controls. We observe a discernable microbial transition from areas of normal skin to the skin surrounding blisters and erosions. Moreover, we find a loss of taxa associated with protective immunity functions and a relative increase in *Staphylococcus aureus*, a known inflammation- promoting species, in patients with BP. Interestingly, these microbial changes are ubiquitously associated with disease status, regardless of sampling site and the presence of blisters. In anorexia nervosa, we find significant associations between the highly abundant skin antimicrobial peptide psoriasin and features of the skin microbiota, including Shannon diversity as well as the abundance of *Abiotrophia*, an indicator for the healthy-weight control group. Moreover, we observe a significant correlation between body mass index and *Lactobacillus* abundance, another indicator for the healthy-weight group. Collectively, these data suggest that that skin microbiota may play an important role in the emergence of inflammatory disease, perhaps via the loss of beneficial taxa on the skin.

The role of the microbiome in personalized medicine

There are clinically relevant applications of microbiome research that might inform the development of future therapies and public health recommendations. One potential clinical application is autologous skin microbiota transplantation in the treatment of inflammatory skin lesions, as demonstrated by Nakatsuji and colleagues (2017, 2021) in atopic dermatitis. Accordingly, microbial strains that exhibit proinflammatory and anti-inflammatory activities could be isolated from skin lesions. Then, *in vitro* testing of pathogen susceptibility to bacteriocins and microbiota strain(s) that produce bacteriocins could be determined on an individual basis. Such data would advance our understanding of strain-level variation on host-microbe interactions in patients with inflammatory skin disease. Moreover, clinical trials could be designed to study the application of strains identified as having anti-inflammatory activity (i.e., transplanted) onto lesional skin.

Our observations of skin microbiota in BP parallel those previously described in the inflammatory skin disease atopic dermatitis (AD). Atopic and healthy skin differ in bacterial colonization, with atopic lesions harboring greater amounts of potentially pathogenic bacteria like *S. aureus* (Kong et al., 2012; Carmona-Cruz et al., 2022). However, in AD, it seems that the identification of protective versus harmful *Staphylococcus* strains is not unambiguous and may depend on the specific skin microbiota community members. For example, it has been reported that *S. epidermidis* abundance in AD correlates with disease severity and that *S. epidermidis* can cause pro-inflammatory proteolytic barrier-disturbing changes (Cau et al., 2021). Recent work by Rademacher and colleagues (2019) demonstrates that AD-derived *S. epidermidis* strains exhibit pro-inflammatory action in a 3D skin model.

Bacterial traits may contribute to non-infectious skin pathologies, including inflammatory disorders such as psoriasis (Kong, 2011; Carmona-Cruz et al., 2022), atopic dermatitis (Grice and Segre, 2011; Kong et al., 2012; Carmona-Cruz et al., 2022), and autoimmune skin blistering diseases (Srinivas et al., 2013; Belheouane et al., 2022). In BP, we observe strong signals for *Staphylococcus aureus* and evidence of antagonistic interactions between *S. aureus* and other *Staphylococcus* species, including *S. epidermidis* and *S. hominis* (Belheouane et al., 2022). In AD, it seems *Staphylococcus* strains usually presumed to be protective, may in fact be contributing to the disease state (Cau et al., 2021; Rademacher et al., 2019). These findings underscore the importance of a personalized approach for the identification and application of appropriate microbiota to achieve a beneficial, individualized therapeutic response. In doing so,

clinicians could better identify pathogenic strains and vulnerable patients based on microbiota community structure, thereby improving both the diagnosis and prognosis of inflammatory skin diseases.

So far, microbiome-derived therapies have been limited to a few exemplary diseases, including Crohn's disease, ulcerative colitis, and recurrent *Clostridioides difficile* infection (Yadav and Chauhan, 2022). For these diseases, clinical data suggest that microbial dysbiosis can be restored to healthy conditions by adding beneficial microbial strains, removing disease- causing pathogens, and/or manipulating host-microbe interactions (Marchesi et al., 2016; Mimee et al., 2016; Yadav and Chauhan, 2022). Moreover, using an individualized approach to develop targeted antibiotics could help to prevent and control the spread of antibiotic resistance when treating skin wounds with broad-spectrum antibiotics.

Improving reproducibility and reliability in microbiome research

Improvements in technology, techniques, and analysis methods have led us to re-evaluate previous findings in human microbiome research. The previously held "sterile-womb" and "sterile-lungs" beliefs represent how improvements in microbiome study design and workflows can lead to shifting paradigms in medicine. Early work demonstrating the presence of microbiota in the uterus and associated tissues were plagued with concerns that contamination was the driving source of identified bacterial content (Agostinis et al., 2019; Stinson et al., 2019). However, through improved workflows aimed at decreasing the impact of potential contaminants, microbiome researchers recently observed that the uterus and associated tissues likely do harbor their own bacteria (Stinson et al., 2019). Likewise, it was long accepted medical doctrine that the lungs were sterile. Stringent microbiome research now indicates microbiota inhabit healthy lungs and researchers have uncovered a diverse and complex microbial-lung ecosystem that contributes to critical aspects of host biology and development (Hilty et al., 2010; Charlson et al., 2011; Dickson and Huffnagle, 2015; Dickson et al., 2016).

In our effort to further characterize the association between *Nell2* and the candidate bacterium unclassified *Betaproteobacteria*, we find that the impact of potential contaminants might havse obscured biologically relevant associations. We were unable to replicate findings that skin microbiota significantly varies according to the examined genotype. Nevertheless, these negative findings are similarly important to the advancement of science. Reporting negative results can help future scientists better focus their research plans to increase the likelihood of success and can help researchers save precious time and money. Crucially, the publication of negative results strengthens the checks and balances system so that science can "self-correct" with evolving evidence (Weintraub, 2016; Bespalov et al., 2019). Future research here may continue to focus on the functional analysis of host-microbe and microbe-microbe interactions to understand the relevance of the host genome on skin microbiota structure and assembly. Moreover, our findings are an indication that the use of sample DNA concentrations or bacterial load data are crucial for assessing the impact of contamination.

Final remarks

Disentangling the relationship between host and microbiota within the context of human skin disorders is essential for the development of novel therapeutic approaches using microbes or bacterial products. Managing disease symptoms, such as skin blistering, erosion, or intense itching, through the topical application of a bacterial mixture, for example, would likelyimpart minimal systemic side effects and could greatly increase quality of life for patients. The potential clinical applications derived from skin microbiome research are great, but so are the challenges of reaching this goal. Indeed, so much variation exists between healthy individuals that we have yet to clearly define the features of a "healthy skin microbiome" (Oh et al., 2014; Vandegrift et al., 2017). Even so, the potential payoffs are worth aiming high. The characterization of skin

microbiota in inflammatory skin diseases is a crucial first step for gaining novel insights into how microbes interact with each other and their hosts.

Understanding how microbes and host environment, diet, physiology, and genetics intersect to assemble the microbiome might one day lead to sustainable, specialized treatments for treating or preventing disease.

References

Agostinis C, Mangogna A, Bossi F, Ricci G, Kishore U, Bulla R. Uterine Immunity and Microbiota: A Shifting Paradigm. Front Immunol. 2019;10:2387. https://doi.org/10.3389/fimmu.2019.02387.

Belheouane M, Hermes BM, Van Beek N, Benoit S, Bernard P, Drenovska K, et al. Characterization of the skin microbiota in bullous pemphigoid patients and controls reveals novel microbial indicators of disease. J Adv Res. 2022. https://doi.org/10.1016/j.jare.2022.03.019.

Bespalov A, Steckler T, Skolnick P. Be positive about negatives–recommendations for the publication of negative (or null) results. Eur Neuropsychopharmacol. 2019;29:1312–20. https://doi.org/10.1016/j.euroneuro.2019.10.007.

Carmona-Cruz S, Orozco-Covarrubias L, Sáez-de-Ocariz M. The Human Skin Microbiome in Selected Cutaneous Diseases. Front Cell Infect Microbiol. 2022;12.

Cau L, Williams MR, Butcher AM, Nakatsuji T, Kavanaugh JS, Cheng JY, et al. Staphylococcus epidermidis protease EcpA can be a deleterious component of the skin microbiome in atopic dermatitis. J Allergy Clin.
Immunol 2021;147:955-966.e16. https://doi.org/10.1016/j.jaci.2020.06.024.

Charlson ES, Bittinger K, Haas AR, Fitzgerald AS, Frank I, Yadav A, et al. Topographical continuity of bacterial populations in the healthy human respiratory tract. Am J Respir Crit Care Med. 2011;184:957–63. https://doi.org/10.1164/rccm.201104-0655OC.

Dickson RP, Erb-Downward JR, Martinez FJ, Huffnagle GB. The Microbiome and the Respiratory Tract. Annu Rev Physiol. 2016;78:481–504. https://doi.org/10.1146/annurev-physiol-021115-105238.

Dickson RP, Huffnagle GB. The Lung Microbiome: New Principles for Respiratory Bacteriology in Health and Disease. PLoS Pathog. 2015;11:e1004923. https://doi.org/10.1371/journal.ppat.1004923.

Grice EA, Segre JA. The skin microbiome. Nat Rev Microbiol. 2011;9:244–53. https://doi.org/10.1038/nrmicro2537.

Hilty M, Burke C, Pedro H, Cardenas P, Bush A, Bossley C, et al. Disordered Microbial Communities in Asthmatic Airways. PLoS One. 2010;5:e8578. https://doi.org/10.1371/journal.pone.0008578.

Kong HH. Skin microbiome: genomics-based insights into the diversity and role of skin microbes. Trends Mol Med. 2011;17:320–8. https://doi.org/10.1016/j.molmed.2011.01.013.

Kong HH, Oh J, Deming C, Conlan S, Grice EA, Beatson MA, et al. Temporal shifts in the skin microbiome associated with disease flares and treatment in children with atopic dermatitis. Genome Res. 2012;22:850–9. https://doi.org/10.1101/gr.131029.111.

Marchesi JR, Adams DH, Fava F, Hermes GDA, Hirschfield GM, Hold G, et al. The gut microbiota and host health: a new clinical frontier. Gut. 2016;65:330–9. https://doi.org/10.1136/gutjnl-2015-309990.

Mimee M, Citorik RJ, Lu TK. Microbiome therapeutics - Advances and challenges. Adv Drug Deliv Rev. 2016;105:44–54. https://doi.org/10.1016/j.addr.2016.04.032.

Nakatsuji T, Chen TH, Narala S, Chun KA, Two AM, Yun T, et al. Antimicrobials from human skin commensal bacteria protect against Staphylococcus aureus and are deficient in atopic dermatitis. Sci Transl Med. 2017;9. https://doi.org/10.1126/scitranslmed.aah4680.

Nakatsuji T, Hata TR, Tong Y, Cheng JY, Shafiq F, Butcher AM, et al. Development of a human skin commensal microbe for bacteriotherapy of atopic dermatitis and use in a phase 1 randomized clinical trial. Nat Med. 2021:1–10. https://doi.org/10.1038/s41591-021-01256-2.

Oh J, Byrd AL, Deming C, Conlan S, Kong HH, Segre JA. Biogeography and individuality shape function in the human skin metagenome. Nature 2014;514:59–64. https://doi.org/10.1038/nature13786.

Rademacher F, Simanski M, Hesse B, Dombrowsky G, Vent N, Gläser R, et al. Staphylococcus epidermidis Activates Aryl Hydrocarbon Receptor Signaling in Human Keratinocytes: Implications for Cutaneous Defense. J Innate Immun. 2019;11:125–35. https://doi.org/10.1159/000492162.

Srinivas G, Möller S, Wang J, Künzel S, Zillikens D, Baines JF, et al. Genome-wide mapping of gene– microbiota interactions in susceptibility to autoimmune skin blistering. Nat Commun. 2013;4:1–7. https://doi.org/10.1038/ncomms3462.

Stinson LF, Boyce MC, Payne MS, Keelan JA. The Not-so-Sterile Womb: Evidence That the Human Fetus Is Exposed to Bacteria Prior to Birth. Frontiers in Microbiology. 2019;10.

Vandegrift R, Bateman AC, Siemens KN, Nguyen M, Wilson HE, Green JL, et al. Cleanliness in context: reconciling hygiene with a modern microbial perspective. Microbiome. 2017;5:76. https://doi.org/10.1186/s40168-017-0294-2.

Weintraub PG. The Importance of Publishing Negative Results. J Insect Sci. 2016;16:109. https://doi.org/10.1093/jisesa/iew092.

Yadav M, Chauhan NS. Microbiome therapeutics: exploring the present scenario and challenges. Gastroenterology Report. 2022;10:goab046. https://doi.org/10.1093/gastro/goab046.

Acknowledgements

I pivoted towards a scientific career in 2015 when I made an emotional and life-changing decision to leave my livelihood as an alternative medicine practitioner. From 2011 to 2014, I worked as a naturopathic doctor ("Heilpraktiker") in the United States. However, I discovered that natural medicine equates with unproven, unethical, and often dangerous practices. A first- hand experience with a colleague prescribing an illegal drug to cancer patients epitomizes the immense harm that can arise when "natural health" practitioners attempt medical care. This watershed moment manifested as a personal crisis—I had to reconsider everything I thought to be true about medicine and science. I reported this colleague to the authorities and decided to leave my career. One year later, I entered a Master of Science program in at Kiel University to re-educate myself in science and move decisively into a credible field.

In 2017, I accepted a doctoral position within the research groups of Prof. Dr. John Baines and Prof. Dr. Saleh Ibrahim to research the mammalian skin microbiome and its complex interactions with host genetics. Over the last five years, I have found within myself the tenacity for problem-solving, scientific questioning, and hard work. I have become a wife, a mom, a science communicator, and now, a scientist. Indeed, I contain multitudes.

First, I thank my advisors Dr. John Baines and Dr. Saleh Ibrahim, whose steadfast support bolstered me through some of my greatest challenges. Dr. Baines embodied the role as my "Doktorvater" in every sense of the word. For teaching, mentoring, and guiding me with kindness and understanding throughout my research, I thank you. I appreciate our phone calls that often drifted into discussing American pop culture and comedy. These mini-mental breaks helped lessen home sickness and provided much needed respites. Mostly, though, I will forever be grateful to you for not only supporting me these past years, but for also supporting my family. I cannot find the words to express how reassuring it was to know that you have always been on my side.

Dr. Ibrahim provided enthusiasm for this body of work, encouragement, and support as a confidant in moments plagued by stress and worry. I'll never forget Dr. Ibrahim's response when I told him I was pregnant, "Oh, how wonderful. This is such good news." This show of kindheartedness and genuine happiness, despite the obvious delays and challenges it would introduce, was overwhelming. I cherish this memory.

I thank the DFG and Research Training Group (RTG) Genes, Environment, and Inflammation in conjunction with Kiel University and the University of Lübeck for funding this work.

I am indebted to many talented members of the global scientific and skeptical community. Prof. Dr. Edzard Ernst and Jann Bellamy, JD opened the door for my career change. When I first emailed each of them, telling my story of loss and confusion around leaving the profession of naturopathy, they responded with compassion and empathy. Thank you for helping me become a better person.

I have been fortunate to garner the support of many individuals in the skeptical community. Eran Segev spearheaded a worldwide fundraising effort on my behalf when I was sued for libel in Germany after writing critical articles about an American naturopath who uses dangerous therapies to treat cancer. He handled all the logistics involved with setting up such a fund and managed the legal payments. In doing so, he unburdened me from a massive amount of emotional stress, which allowed me to continue to focus on my doctoral research. Eran, thank you, my dear friend, for helping me at such a trying time.

Many other friends and organizations were involved in the logistics and success of this fundraising effort and provided me with support, including the Australian Skeptics Society, Marsh (Michael Marshall) and Dr. Simon Singh of The Good Thinking Society, the European Council of Skeptical Organizations, Dr. Steven Novella, Dr. David Gorski, Science-Based Medicine, Jonathan Jerry and Dr. Joe Schwarcz, Sense about Science, the Center for Inquiry, and more. The out-pouring of love and support from scientists and skeptics around the world was incredible and truly sustaining, even in the worst moments. Thank you, all. I am truly grateful.

In 2018, I was awarded the prestigious John Maddox Prize by *Nature* and Sense about Science in honor of standing up for science in the face of adversity. I was nominated for this award by Dr. Chris French from Goldsmiths, University of London, and Dr. Chris Peters from Sense about Science with additional recommendations from Dr. Edzard Ernst and Dr. Simon Singh. This special recognition came only five months after giving birth to my daughter and during the initial development of my doctoral projects. I thank you all for recognizing the hardship I faced. Your support motivated me to keep writing and sharing my experiences with others.

I am deeply grateful for the support of my friends, and especially those I turned to for counsel mid-way through my doctoral work. Thank you, Dr. Kevin Folta, Cara Santa Maria, and Dr. Thomas Mohr, for each spending the time talking with me about my research and encouraging me to carry on. Thank you, Robert Starr, for reminding me to keep balance, when possible, and for our meandering discussions about life.

I thank the members of the Baines and Ibrahim working groups for being my sounding board and for putting up with all my complaining. Dr. Shauni Doms, you are my "lab bestie" and my real life bestie. You helped me in so many ways, from keeping me company in the lab to sewing clothes for Kira—I love you and I am so grateful you are in my life. Thank you to Dr. Malte Rühlemann for reading countless lines of code and patiently explaining the ins and outs of so many analyses. You made my research better. Thank you to Dr. Nadia Andreani for teaching me and supporting me in the lab, even when that meant waking up at 5 AM or coming to the lab on a weekend. Your optimism and hard work helped me reach the finish line. To Katja Cloppenborg-Schmidt and Yasmin Claussen, thank you for being amazing lab technicians and for helping me with lab work. Dr. Misa Hirose and Adina-Malin Tietje each helped me get through the part of my dissertation that I most dreaded: the mouse work. I thank each of you for the long days in the dissection room and countless hours spent organizing the *Nell2* project.

My family has provided me with support and love at every turn. Thank you to my parents, Timothy and Patti Deegan, for rooting for my success in all that I attempt. Dad, your advice to "show up every day," has served me so well and gotten me through hard times. When I wanted to quit, I remembered your advice to just show up. So, I did, and eventually, I found my way. Thank you. Mom, I get my "grit" from you and those farm-girl genes. I am lucky you are mine. Thank you to my siblings, Matthew and Sara Deegan, for your friendship and love. Matthew, you're the best fun uncle; Thank you for making wonder, magic, and play central to our household. Sara, thank you for your encouragement and the gift of sisterly love. There is nothing better.

Thank you to the best dog in the whole word, Buddha, for reminding me to stand up from the computer desk and take a walk. You're a good boy.

On June 23, 2018, my daughter, Kira Jane, was born healthy with a feisty spirit that has kept me going (and kept me awake) when I did not think I had the energy to continue. You, my sweet girl, have perhaps sacrificed the most during my dissertation. You have tolerated my Kiel work trips, long working days, skipped trips to the park, zoo, and parties, and cancelled playdates, so that I could complete this thesis. Thank you, little one, for the love, kisses, and cuddles. Let's take that

beach trip now.

My husband has provided unwavering support and love throughout my studies and science activism, even when his own career and emotional well-being were threatened. Thank you, Dr. Taylor Hermes, for showing me unconditional love, modeling patience, taking on more than your fair share of childcare duties for long stretches, and prioritizing my needs over yours. We've made a wonderful life together, full of love, laughter, and ggplots. You and Kira are everything. This is for us.