2015

Novel Insights into Skin Bacterial and Viral Communities in Health and Acute Wounding

Geoffrey Daniel Hannigan
University of Pennsylvania, ghanni@upenn.edu

Novel Insights into Skin Bacterial and Viral Communities in Health and Acute Wounding

Abstract

Human skin is colonized by diverse microbial communities that have broad impacts on health and disease. Bacterial communities have been associated with dermatological diseases including Atopic Dermatitis and Psoriasis, and while roles of virus communities (viromes) in cutaneous health are poorly characterized, virome dysbiosis has been implicated in other human diseases and individual viruses are known to impact cutaneous health. Here we present a comprehensive research program aimed at broadly understanding the roles of bacteria and viruses in human dermatological health and perturbation by wounding. In the first section, we characterize the healthy human skin virome and investigate potential interactions between virus and bacterial communities. Samples were collected from sixteen subjects at eight body sites over one month. Virome diversity and composition varied by natural skin occlusion and the microenvironment substrates. Viruses were enriched for temperate replication-cycle genes, and maintained genes encoding potential antibiotic resistance and virulence factors. We also highlighted potential interactions between the virus (phage) and bacterial communities, including CRISPR targeting and significant ecological associations by co-occurrence modeling. This work provides a greater ecological context for our individualized understanding of cutaneous viruses, and provides a foundation for future studies of the skin virome upon perturbation and disease. In the second section, we characterize the microbial communities associated with skin perturbation in the form of acute, open fracture wounding. Thirty subjects presenting to the Hospital of the University of Pennsylvania for acute care of open fractures were enrolled in a prospective cohort study. Microbiota were collected from wound center and adjacent skin upon presentation to the ER and during follow up visits. Bacterial communities were studied using 16S rRNA amplicon sequencing. Microbiome composition and diversity colonizing open fracture wounds became increasingly similar to adjacent skin microbiota with healing. Clinical factors were associated with various aspects of microbiota diversity and composition. We also developed an analysis tool patPRO to facilitate analysis of this longitudinal dataset, and to aid others in analyses of similar data. The results of this pilot study demonstrate the diversity and dynamism of the open fracture microbiota, and their relationship to clinical variables.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Cell & Molecular Biology

First Advisor

Elizabeth A. Grice

Keywords

Infection, Microbiome, Open Fracture, Virome

Subject Categories

Microbiology

NOVEL INSIGHTS INTO SKIN BACTERIAL AND VIRAL COMMUNITIES IN HEALTH AND ACUTE WOUNDING

Geoffrey Daniel Hannigan

A DISSERTATION

in

Cell & Molecular Biology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2015

Supervisor of Dissertation

_____

Elizabeth A. Grice PhD.
Assistant Professor of Dermatology

Graduate Group Chairperson

_____

Daniel S. Kessler, PhD.
Associate Professor of Cell and Developmental Biology

Dissertation Committee

Frederic D. Bushman, PhD.
Professor of Microbiology

Jean Bennett, M.D., PhD.
Professor of Ophthalmology

David B. Weiner, PhD.
Professor of Pathology and Laboratory Medicine

John R. Stanley, M.D.
Professor of Dermatology

NOVEL INSIGHTS INTO SKIN BACTERIAL AND VIRAL COMMUNITIES IN HEALTH AND

ACUTE WOUNDING

DEDICATION

For my wife Brittany, my brother Christopher, and my parents Daniel and JoAnn.

ACKNOWLEDGMENTS

I am incredibly blessed to have worked with, and to have been advised by, absolutely outstanding people. One of the most influential mentors I have had is my graduate research advisor, Elizabeth Grice. Elizabeth has been an influential mentor, and has continued to support my growth as a scientist and as a person. Her kindness, passion for excellence, and support for my growing independence as a scientist has shaped me to be a better researcher, teacher, and overall person. I count myself as incredibly fortunate to have trained under Elizabeth, and she will continue to be a role model as I continue in my scientific career.

For their guidance and support, both personal and scientific, I thank the members of my thesis committee. Rick Bushman has been an amazing mentor and advocate throughout my pursuits in microbiome research and the virome. Dave Weiner is an incredible leader, teacher, and student advocate, and I cannot thank him enough for his guidance as well. He has helped me get the most out of my graduate school career, and made it an enjoyable experience. Jean Bennett and John Stanley have also been incredible mentors and I cannot thank them enough for everything they have done.

For their continued support and camaraderie, both inside and outside of lab, I am a better person for having known the past and present members of the Grice Lab (Club Grice): Amanda Tyldsley, Jackie Meisel, Brendan Hodkinson, Adam SanMiguel, Michael Loesche, Joseph Horwinski, Qi Zheng, Casey Bartow-McKenny, Georgia Sfyroera, and Ana Misic. I am also appreciative for the members of the Bushman Lab for their continued support. I am thankful to have worked with Alexandra Bryson, Christel Chehoud, Aubrey Bailey, Kyle Bittinger, Jacque Young, and the rest of the lab. I am especially thankful for Sam Minot, who has continued to be a generous source of guidance and advice, even after graduating and moving on from Penn.

I am thankful for the continued support of our Dermatology department, including John Stanley who mentored me as a member of my thesis committee. I am also thankful for the support of Misha Rosenbach, George Cotsarelis, John Seykora, Steve Prouty, and the other members of the University of Pennsylvania Department of Dermatology.

My scientific training began before I attended graduate school, and I was fortunate to have had so many great mentors. Dave Rowe, my high school science teacher and Science Olympiad coach, was the most influential person in my decision to pursue research in microbiology. I am also thankful for the continued support of Jean Regal and John Dahl, who helped me develop my research skills as I was starting as an undergraduate student.

Finally, none of this would have been possible without the incredible support of my family. My wife Brittany has been a constant source of strength and motivation, and I cannot say enough about how crucial she has been in my life. I could not have accomplished any of this without her. My parents JoAnn and Daniel have never ceased supporting me, and this would not have been possible without them. They have been amazing parents and role models, and I am incredibly fortunate. My brother Christopher has been with me since before I can remember, and has continued to push me to do my best in all my endeavors. I am also appreciative for the support of my family-in-law, including Jean and Martin Bauer. I am especially appreciative of Ashley Bauer, who has been a great colleague and friend as we pursue scientific careers. Finally, I thank God for surrounding me with so many wonderful people and opportunities.

ABSTRACT

NOVEL INSIGHTS INTO SKIN BACTERIAL AND VIRAL COMMUNITIES IN HEALTH AND ACUTE WOUNDING

Geoffrey Daniel Hannigan

Elizabeth A Grice, PhD

Human skin is colonized by diverse microbial communities that have broad impacts on health and disease. Bacterial communities have been associated with dermatological diseases including Atopic Dermatitis and Psoriasis, and while roles of virus communities (viromes) in cutaneous health are poorly characterized, virome dysbiosis has been implicated in other human diseases and individual viruses are known to impact cutaneous health. Here we present a comprehensive research program aimed at broadly understanding the roles of bacteria and viruses in human dermatological health and perturbation by wounding. In the first section, we characterize the healthy human skin virome and investigate potential interactions between virus and bacterial communities. Samples were collected from sixteen subjects at eight body sites over one month. Virome diversity and composition varied by natural skin occlusion and the microenvironment substrates. Viruses were enriched for temperate replication-cycle genes, and maintained genes encoding potential antibiotic resistance and virulence factors. We also highlighted potential interactions between the virus (phage) and bacterial communities, including CRISPR targeting and significant ecological associations by co-occurrence modeling. This work provides a greater ecological context for our individualized understanding of cutaneous viruses, and provides a foundation for future studies of the skin virome upon perturbation and disease. In the second section, we characterize the microbial communities associated with skin perturbation in the form of acute, open fracture wounding. Thirty subjects presenting to the Hospital of the University of

Pennsylvania for acute care of open fractures were enrolled in a prospective cohort study. Microbiota were collected from wound center and adjacent skin upon presentation to the ER and during follow up visits. Bacterial communities were studied using 16S rRNA amplicon sequencing. Microbiome composition and diversity colonizing open fracture wounds became increasingly similar to adjacent skin microbiota with healing. Clinical factors were associated with various aspects of microbiota diversity and composition. We also developed an analysis tool patPRO to facilitate analysis of this longitudinal dataset, and to aid others in analyses of similar data. The results of this pilot study demonstrate the diversity and dynamism of the open fracture microbiota, and their relationship to clinical variables.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 - A Primer to Pursuing Broad Studies of the Skin Microbiome and Virome

1.1 Introduction

The human body is colonized by a variety of microbes collectively referred to as the microbiome. These microbes play diverse roles in human health and disease across various human systems, including the gut, oral cavity, vagina, and the skin [1, 2]. The most studied of these microbial communities are bacteria, whose dysbiotic states are thought to promote a broad range of diseases including Clostridium difficile infections [3], atoptic dermatitis [1], and even chronic wounds [4, 5]. Individual viruses, including bacteriophages (bacterial viruses), are also known to impact human health, but the impacts of human virus communities as a whole on health and disease are less understood.

There are diverse viruses associated with the skin, and these individual viruses have broad impacts on human cutaneous health. Human viruses are know to directly impact human health through infection of the human cells. Examples of these interactions include Human Papillomavirus infections which can cause cutaneous warts, and the Merkel Cell Polyomavirus which can promote merkel cell carcinoma, a very serious and aggressive form of skin cancer [1].

Viruses can also indirectly impact human health, most often as bacteriophages that can modulate bacterial community composition and gene expression, resulting in an effect on the health of the human substrate. Phages can impact bacterial gene expression through various mechanisms including transduction, which is the process by which phages transport genes throughout bacterial communities.

While we know a lot about the effects of individual viruses on the skin, we know a lot less about the virus communities associated with the skin. Although studies of the human virome remain in their infancy, virus dysbiosis has been associated with human disease, including periodontal disease [6]. Bacteriophages are also known to promote horizontal gene transfer through transduction, including transfer of antibiotic resistance genes, but this has not yet been evaluated in a skin community setting [7].

Previous studies have begun to investigate the skin virome, but have only studied them as subsets of a larger whole skin metagenome (including sequences of the human host, bacteria, fungi, and viruses) [8, 9]. Because human, bacterial, and fungal genomes are orders of magnitude larger than most viruses, and because the longer genomes in the sample are more likely to be sequenced, these approaches result in low virome coverage. These studies are also less equipped to evaluate the virus dark matter (unknown and uncharacterized members) within the communities because the majority of these methods rely on purified virus sample sets.

1.2 Advancing Our Understanding of the Microbiome in Health: The Human Skin Virome

To address these knowledge gaps and establish a foundation on which further studies can be built, the first section of this work focuses on characterizing the healthy human skin virome and its potential interactions with their associated bacterial communities. This study is powered over previous studies by its use of optimized virus purification techniques that do not require us to study the virome as a subset of the whole metagenome. This study is also powerful because of the robust and sophisticated analysis methods implemented, and the depth of the analysis.

1.3 Advancing Our Understanding of the Microbiome in Disease: The Open Fracture Wound Microbiome

While we furthered the field's understanding of the healthy skin microbiome by characterizing its viral component, we also studied the microbial communities associated with disrupted health. More specifically, we investigated open fracture wounds. These wounds involve breaking of the bone through the skin and soft tissue, which puts the patient at an increased risk for infection-related complications. While there is a clear need to understand the microbial communities associated with such a vulnerable state, contemporary microbiology methods had not been used to study such a system. To address this unmet need, we implemented modern, sequence-based microbiological techniques to gain a more robust understanding of the microbial communities associated with these at-risk acute wounds.

This study provides initial insights into the microbial community dynamics throughout the healing process. We also highlight associations between the bacterial communities and important clinical factors. Together these highlight the importance of understanding the microbiome associated with acute traumatic wounds, and lays a foundation for future, more robust studies powered by greater sample sizes. We further present an analysis toolkit for processing and visualizing this longitudinal dataset, and related datasets.

1.4 Summary of the Work to be Presented

Together, this body of work represents a comprehensive contribution to the field of dermatological microbiology. We begin the first section with a detailed introduction to the current state of the skin microbiome, the historical context of the field, the current understanding of health and diseased states, and how microbiome studies will influence therapeutic design. We then present strong research contributions to the overall understanding of the skin virome in health, thus laying a foundation for future studies that wish to investigate the viral communities in diseased or otherwise perturbed states.

We begin the second section of this work with an introduction to our current understanding of open fracture wound care, current best practices for treatment, and the roles of microbes in this type of wound. We then present our study investigating the microbiome dynamics associated with the healing and clinical factors of those wounds. We conclude with a presentation of our analysis toolkit that facilitates visualization of longitudinal studies such as this.

Together this work sheds new light on previously understudied aspects of the human microbiome, including the interactions between virus and bacterial communities. This lays a foundation for future studies of bacterial and viral communities, especially related to acute wounds.

1.5 Works Cited

1. Hannigan GD, Grice EA: **Microbial ecology of the skin in the era of metagenomics and molecular microbiology.** *Cold Spring Harb Perspect Med* 2013, **3:**a015362.
2. Grice EA, Segre JA: **The human microbiome: our second genome.** *Annu Rev Genomics Hum Genet* 2012, **13:**151-170.
3. Schubert AM, Sinani H, Schloss PD: **Antibiotic-Induced Alterations of the Murine Gut Microbiota and Subsequent Effects on Colonization Resistance against Clostridium difficile.** *MBio* 2015, **6:**e00974.
4. Hannigan GD, Hodkinson BP, McGinnis K, Tyldsley AS, Anari JB, Horan AD, Grice EA, Mehta S: **Culture-independent pilot study of microbiota colonizing open fractures and association with severity, mechanism, location, and complication from presentation to early outpatient follow-up.** *J Orthop Res* 2014, **32:**597-605.
5. Hannigan GD, Pulos N, Grice EA, Mehta S: **Current Concepts and Ongoing Research in the Prevention and Treatment of Open Fracture Infections.** *Adv Wound Care (New Rochelle)* 2015, **4:**59-74.
6. Ly M, Abeles SR, Boehm TK, Robles-Sikisaka R, Naidu M, Santiago-Rodriguez T, Pride DT: **Altered oral viral ecology in association with periodontal disease.** *MBio* 2014, **5:**e01133-01114.
7. Balcazar JL: **Bacteriophages as vehicles for antibiotic resistance genes in the environment.** *PLoS Pathog* 2014, **10:**e1004219.
8. Oh J, Byrd AL, Deming C, Conlan S, Program NCS, Kong HH, Segre JA: **Biogeography and individuality shape function in the human skin metagenome.** *Nature* 2014, **514:**59-64.
9. Foulongne V, Sauvage V, Hebert C, Dereure O, Cheval J, Gouilh MA, Pariente K, Segondy M, Burguiere A, Manuguerra JC, et al: **Human skin microbiota: high**

diversity of DNA viruses identified on the human skin by high throughput sequencing. *PLoS One* 2012, **7:**e38499.

CHAPTER 2 – Microbial Ecology of the Skin in the Era of Metagenomics and Molecular Microbiology

The contents of this chapter have been published as:

> Hannigan GD, Grice EA. (2013) Microbial Ecology of the Skin in the Era of Metagenomics and Molecular Microbiology. Cold Spring Harb Perspect Med, Fiona Watt and Tony Oro, Editors. doi: 10.1101/cshperspect.a015362.

2.1 Abstract

The skin is the primary physical barrier between the body and the external environment and is also a substrate for the colonization of numerous microbes. Previously, dermatological microbiology research was dominated by culture-based techniques, but significant advances in genomic technologies have enabled the development of less-biased, culture-independent approaches to characterize skin microbial communities. These molecular microbiology approaches illustrate the great diversity of microbiota colonizing the skin and highlight unique features such as site specificity, temporal dynamics, and interpersonal variation. Disruptions in skin commensal microbiota are associated with the progression of many dermatological diseases. A greater understanding of how skin microbes interact with each other and with their host, and how we can therapeutically manipulate those interactions, will provide powerful tools for treating and preventing dermatological disease.

2.2 Introduction

The skin acts not just as a protective physical barrier between the body and the external environment; it is itself an environmental substrate, harboring a rich and diverse community of microorganisms (the microbiome). The human microbiome includes the bacteria, fungi, viruses,

archaea, and microeukaryotes that inhabit the various body environments, such as the gut, oral cavity, and skin [1]. In recent years, it has become increasingly apparent that the microbiome interacts extensively with the human body and plays roles in immune system development and function, disease etiology and pathology, cancer development, and defense against pathogens.

The skin is a complex ecosystem that maintains topographically distinct microbial populations, as well as distinct environmental niches. Overall, the surface of the skin is cooler than the core body temperature, is slightly acidic, and squames are continuously shed from the skin surface as a result of terminal differentiation [2]. These attributes undoubtedly select for specific microbiota adapted to these unique conditions. The geography of the skin includes sebaceous areas (including face and back), moist areas (including toe/finger web space and arm pit), dry areas (including forearm and buttock), and sites containing varied densities of hair follicles, skin folds, and skin thicknesses. A critical function of the skin microbiota is "colonization resistance," in which commensal microbiota occupy these distinct niches to block colonization and/or invasion by opportunistic or pathogenic organisms.

Community composition equilibrium across the varied geography of the skin is maintained by nutrient and space competition among microbes, production of antimicrobial peptides (AMPs) by commensal microbes and host cells, and modulation of the host immune response by commensal microbes [3]. For example, the skin commensal Staphylococcus epidermidis has been reported to modulate the innate immune response by inhibiting skin inflammation through pattern-recognition receptor-mediated cross talk [4]. Complement, an evolutionarily conserved arm of the innate immune system, was shown to maintain diversity of the skin microbiota in a mouse model, and, conversely, the skin microbiota regulated complement at the gene expression level [5]. The microbiome is also fundamental in adaptive immune system equilibrium at the skin, and skin T-

cell function and the local inflammatory milieu appear to be autonomously controlled by the commensal skin microbiota [6]. These and other findings have contributed to the mounting evidence suggesting that the commensal skin microbiota is intricately involved in both innate and adaptive skin immunity.

For these reasons, a thorough understanding of the commensal skin microbiota is required to gain insight into microbial involvement in skin health and disease. The beneficial role of skin commensals and the pathogenic role of those microbes that cause disease have long been a focus of studies examining the microbial ecology of the skin. Contemporary, culture-independent methods for identifying and characterizing microbial communities have accelerated and added precision to our understanding of host–microbe interactions at the skin surface. In this article, we provide a comprehensive discussion of the human skin microbiome in health and disease states and how this understanding is informing skin disease diagnosis and treatment.

2.3 Dawning of the molecular biology era

The study of the human cutaneous microbiota has a rich history that spans more than five decades [7]. Early methods for studying skin-associated bacteria, fungi, and viruses were limited to culturing the microorganism and defining its phylogeny and taxonomy through phenotypic, microscopic, and biochemical relationships. Dependency on generation of pure cultures introduces inherent biases because the procedures select for the most abundant and rapidly growing microorganisms of the community. Culture-based studies of viruses (including bacteriophage) are further limited because they require coculturing with their prokaryotic or eukaryotic hosts, additionally preventing the identification of viruses associated with unknown hosts. Viruses are also not readily visible by basic microscopic methods and thus are very difficult to characterize by direct morphological observation. Although great insight into

microbial colonization of cutaneous surfaces in health and disease was gained using culture-dependent methods, there were significant limitations to the conclusions that could be drawn.

Advances in DNA sequencing technology and culture-independent methods of microbial identification have greatly enabled high-throughput, detailed characterization of microbial communities. These methods are based on surveys of marker genes, generally conserved, universal genes found in all organisms within particular taxonomic levels. Bacterial communities are most commonly classified by the sequence of their small subunit 16S ribosomal RNA (rRNA) gene (Fig. 1) [8]. These genes contain both conserved regions, which allow for PCR primer binding and phylogenetic analysis, as well as variable regions, whose sequences allow for taxonomic classification. Following amplification and sequencing of 16S rRNA genes, sequence data can be analyzed in a variety of ways, including assignment of taxonomy, phylogenetic analysis, and community analyses (Fig. 1). Fungi are often classified by sequencing of the internal transcribed spacer (ITS) region that lies between the small and large subunit rRNA genes in eukaryotes [9]. In addition to offering clearer definitions for determining microbial taxonomy, conserved gene sequence analysis does not require the microorganism to be cultured, which therefore eliminates those biases associated with culturing procedures.

Unlike bacteria and fungi, viruses and bacteriophage present a special difficulty because they do not contain a consensus gene that can be used for widespread taxonomic identification. Closely related groups may be phylogenetically analyzed using specific conserved genes, such as the human papillomavirus L1 gene, but this is far less robust than the 16S rRNA and ITS sequencing and identification approaches used to classify bacteria and fungi. Comparative genome analysis is complicated by the high frequency of gene transfer among virus and host genomes and the lack of comprehensive, annotated reference databases and assigned taxonomy. A solution to this problem

is the use of whole-genome shotgun metagenomics, which does not rely on amplification and sequencing of marker genes but, rather, allows for sequencing and analysis of the sample's full genetic potential [10]. This type of strategy not only bypasses PCR, but also can provide insight into what microbial communities are doing on the skin surface. These types of approaches are still under development in the skin, because high amounts of host DNA and low amounts of microbial DNA present technical limitations for metagenomic approaches.

2.4 The skin microbiome in health

2.4.1 Bacteria

Before the advent of molecular techniques to characterize skin microbiota, the temporal and topographical diversity of the skin microbiota was still considered vast. Early studies produced variable results in bacterial quantity and taxonomy, hypothesized at the time to be a result of inherent topographical and temporal diversity of skin bacterial communities [11]. In a comprehensive study, which cultured under both aerobic and anaerobic conditions, skin bacterial colonization differed between anatomical sites, and the highest bacterial load was observed in sebaceous sites [11]. Furthermore, skin colonization was dominated by a small group of taxa, including Propionibacterium acnes and Staphylococcus epidermidis. Years later, these same features are apparent using sophisticated sequencing-based techniques to characterize skin microbiota.

Indeed, site-specific colonization is a key feature of the human skin microbiome. Using a 16S rRNA sequencing approach in healthy adults, sebaceous regions were found to be preferentially colonized by Propionibacterium and Staphylococcus spp.; moist sites predominantly maintained Corynebacterium and Staphylococcus spp.; and dry sites, which, despite general variability and greater diversity, displayed a significant presence of β-Proteobacteria and Flavobacteriales [12].

In the same study, 19 bacterial phyla were identified, but skin was dominated by four phyla: the Actinobacteria, Firmicutes, Proteobacteria, and Bacteroidetes (Fig. 2). Another key finding was that longitudinal stability was dependent on the skin site, with sebaceous sites being the most stable and dry sites being the most variable over time. Costello et al. (2009) similarly reported that topographical community variability was greater than temporal variability among individuals. Interestingly, greater microbial diversity characterized skin microbiota, as compared with gut or oral microbiota of the same individuals. Key findings of these and other studies show that skin bacterial communities are generally diverse between individuals [13, 14] and may be influenced by ethnicity, lifestyle, and/or geography, as suggested by a study comparing cutaneous microbiota colonizing South American Amerindians and U.S. residents [15]. Subject sex, handedness, and time since last hand washing also appear to affect bacterial community composition [16]. A subsequent study confirmed the differences in community composition between sexes when investigating the forearm, but found little influence of sex on forehead community composition, thereby supporting early observations of variability among anatomical regions [17].

The human skin microbiota is established immediately after birth, and delivery mode seems to influence the neonate's first skin microbiota. Dominguez-Bello et al. (2010) showed that vaginally delivered neonates were colonized with bacteria similar to those colonizing the mother's vagina (i.e., Lactobacillus, Prevotella, Sneathia spp.), and neonates delivered by Cesarean section were colonized with those bacteria found on the mother's skin (i.e., Propionibacterium, Staphylococcus, Corynebacterium spp.). Studies in infants over the first year of life showed that diversity of skin microbiota increases with age, as does site specificity, and is overall characterized by predominance of the phylum Firmicutes [18]. Upon sexual maturation, the skin becomes colonized by increased amounts of Corynebacterium and Propionibacterium

[19]. Colonization by these lipophilic bacterial taxa is likely a result of hormone-stimulated sebaceous gland activity and increasing sebum production during puberty. Metabolism of the lipids in sebum by these bacterial taxa also decreases the pH of the skin, thus discouraging colonization by other taxa.

The skin was an organ included in the National Institutes of Health Roadmap Human Microbiome Project, in which a cohort of 242 phenotyped healthy adults were subject to sampling of microbiota at various body sites. Their findings confirmed previous smaller-scale studies, by suggesting that the skin microbiota is diverse, but dominated by a small group of genera, in particular Staphylococcus, Propionibacterium, and Corynebacterium [20]. Interestingly, the metabolic and functional pathways encoded by the skin microbiota colonizing the retroauricular crease were more constant and diverse than the taxonomic composition, suggesting low metabolic diversity among a taxonomically diverse population. Furthermore, age was associated with differentially encoded metagenomic pathways on the skin, as well as a decrease in the phylum Firmicutes.

2.4.2 Fungi

Although bacteria represent a major focus of past and present microbiome studies, the fungal microbiota is also thought to play a significant role in skin health and disease. Cultivation-based studies identified the major component of the skin fungal community as Malassezia (formerly known as Pityrosporum) genus, consisting primarily of seven of the 14 known species [21]. These findings have been confirmed by molecular community analysis. Using 18S rRNA gene and ITS region sequencing, Paulino et al. (2006) reported that the skin forearm community is dominated by Malassezia, and further analysis with multiplex real-time PCR (to speciate Malassezia) suggested that the predominant commensal species were Malassezia globosa and Malassezia

restricta, with Malassezia furfur (the dominant species identified by culturing methods) contributing relatively little to the overall community abundance [22]. Conversely, another recent study of three healthy scalps found that Malassezia spp. only account for a small fraction of the commensal fungi on the scalp [23]. A larger-scale, extensive topographical map of the fungal skin microbiota, based on sequencing of the fungal ITS region, confirmed that Malassezia is dominant in most regions of the skin, but sites on the feet (plantar heel, toenail, and toe web) had the greatest fungal diversity of all body sites [24]. Databases and other resources for identifying and analyzing fungal sequences, similar to those used for 16S rRNA gene studies, are still under development, and it is expected that our knowledge of the fungal microbiome will expand as these tools become readily available.

2.4.3 Viruses

One of the most extensively studied human skin viruses is the human papillomavirus (HPV). Although it was originally thought that certain strains were found only in skin cancer lesions, PCR quantification of HPV marker genes revealed that healthy skin is also a habitat for a broad spectrum of HPV strains [25, 26]. Sequence analysis of the conserved L1 open reading frame (ORF) revealed a diverse community of HPV types on healthy skin, while identifying HPVs that were previously unknown [27]. Follow-up studies have confirmed the ubiquity and diversity of HPV types throughout human populations [28-30].

The other major group of commensal human skin viruses is the human polyomaviruses (HPyVs). Polyomaviruses were first described in mice in 1953 but have since been described in numerous animals, including humans [31]. Although originally studied in the context of cancer, they, like the papilloma viruses, have been found on healthy human skin [32]. There are many types of HPyVs, with many only recently discovered using molecular techniques; the most common to

human skin are HPyV6, HPyV7, and Merkel cell polyomavirus (MCHPyV) [31, 32]. A recent study using a whole metagenomic analysis of the human skin virome of healthy and cancerous individuals confirmed a cutaneous viral microbiota dominated by HPVs, HPyVs, and circoviruses [33]. These studies are still in their early stages, and as new virus species continue to be discovered and new analysis strategies developed, future studies will likely continue to characterize the viral community diversities and pathogenic/oncogenic potential.

The other viral component of the microbiome is the bacteriophage, about which little is known in the skin. Recent studies have used culture-dependent techniques paired with genomic analyses, as well as analysis of prophage (bacteriophage integrated into the bacterial host genome), to understand and characterize the genomic diversity of subsets of the skin bacteriophage communities, such as the limited diversity of Propionibacterium bacteriophage and the diversity of Staphylococcus bacteriophage [34-36]. Whole metagenomic shotgun sequence analysis of skin swabs from five healthy patients and one patient with a previous Merkel cell carcinoma lesion indicate that two families dominate cutaneous bacteriophage communities, the Microviridae and Siphoviridae [33]. Further studies will be required to perform more in-depth and functionally informative analysis of the bacteriophage inhabiting the skin, such as characterization of bacterial antibiotic resistance genes maintained in bacteriophage genomes that may be horizontally transferred among bacteria.

Bacteriophages are also known to be important, yet complicated mediators of bacterial horizontal gene transfer through a process known as transduction. Commensal bacteriophage metagenomes have been shown to maintain antibiotic-resistance genes, as well as mediate their transfer between bacteria, in gut and sputum samples from cystic fibrosis patients [37-39]. In the skin, bacteriophage communities have been suggested as mediators of resistance gene transfer between

bacteria [40, 41]. However, more research is required to understand these complicated interactions between bacteria and skin bacteriophage communities.

2.5 The skin microbiome in disease

2.5.1 Atopic Dermatitis

Atopic dermatitis (AD) is a chronic, recurring inflammatory skin disease that occurs more frequently in children than in adults, and has been associated with skin colonization by Staphylococcus aureus. Although no clear microbial cause has been established, antibiotics, corticosteroids, and dilute bleach baths have been relatively effective in the treatment of AD [42]. Furthermore, the enormous increase in incidence over the past three decades with no clear cause raises the interesting possibility that the skin microbiota may modulate gene–environment interactions at the skin surface.

Bacterial virulence factors may in part explain the long-recognized pathogenic association between AD flares and increased colonization by S. aureus. Severe AD development was reported in a mouse model with reduced skin barrier function upon exposure to Staphylococcal protein A (SpA) (Terada et al. 2006). The detection of SpA among 89 children with AD lesions was evaluated as occurring in 91.0% of patients upon presentation, decreasing to 55.6% of patients after antibiotic therapy [43]. Furthermore, there was a significant positive correlation between the levels of SpA and the clinical severity of the lesions. It has also been reported that AD lesions contain increased levels of lipotechoic acid, a known immune-stimulating molecule derived from Gram-positive bacterial cell walls, whose presence further suggests a role of bacterial components in disease [44].

Although S. aureus likely in part contributes to disease pathogenesis, a role for the greater microbial community has recently been investigated. In a 16S-rRNA-based study that analyzed skin microbiota during the course of AD flares and improvement, a correlation between increased disease severity and decreased bacterial diversity was observed, along with altered microbial community structure in AD patients as compared with healthy controls (Fig. 3A) [45]. Bacterial community diversity was also shown to increase after standard AD treatment. Additionally, fungal communities have been shown to change in composition as disease severity progresses [46, 47]. Infants who develop AD maintained early fecal microbiota with less diversity than the early fecal microbiotas of patients who did not develop the disease [48]. Other studies have also shown that altered bacterial compositions of the infant gut microbiota precede the development of AD [49, 50].

Mutations in the gene-encoding filaggrin, an epidermal structural and hydration protein, have been associated with atopic dermatitis and other ichthyotic disorders [51]. Analysis of skin microbiota of a mouse model with a hypomorphic mutation in St14, encoding the serine protease matriptase that regulates filaggrin processing, showed a selective shift in bacterial populations, with increased Corynebacterium and Streptococcus and decreased Pseudomonas species [52]. These findings provide a link between filaggrin deficiency, a common genetic feature of AD, and changes in the skin microbiota.

2.5.2 Psoriasis

Multiple clinical observations support a role for dysbiosis of the skin microbiota in the pathogenesis of psoriasis including the clinical efficacy of topical corticosteroids in the treatment of psoriasis [53] and the observation that a variant of plaque psoriasis, guttate psoriasis, is triggered by Streptococcus infection. Xenograft models of psoriasis, in which unaffected skin

from psoriasis patients was grafted on immunodeficient mice, showed that stimulation with bacterial antigen could trigger the skin to become lesional [54]. Early culture-based studies examining microorganisms associated with psoriasis identified Malassezia, group A and B β-hemolytic streptococci, S. aureus, and Enterococcus faecalis [55-57]. Culture-independent analysis of fungal microbiota found no conclusive evidence to link Malassezia with psoriasis [22, 58]. Analyses of the bacterial microbiota by 16S rRNA gene-based approaches in cross-sectional studies suggest underrepresentation of Propionibacterium and increased representation of the phylum Firmicutes in psoriatic plaques as compared with healthy controls or uninvolved limb skin (Fig. 3B) [59, 60]. Longitudinal studies of the skin microbiota in psoriasis plaques may provide insight into the role of microbes in triggering, propagation, and maintenance of plaques.

2.5.3 Acne Vulgaris

Acne vulgaris is a common skin disorder characterized by abnormalities of sebum production by the pilosebaceous unit, bacterial proliferation, and inflammation. The etiology and pathogenesis of acne remain unclear, but there has been significant evidence supporting microbial roles in the disease. The primary microbe associated with development of acne is Propionibacterium acnes, also a prominent member of the commensal skin microbiota. Topical and systemic antibacterial drugs have long been used to treat acne, with the efficacy commonly attributed to decreased P. acnes colonization and/or activity [61]. Strain-level analysis of the 16S rRNA gene showed that, although the relative abundances of P. acnes did not significantly differ between healthy and acne patients, the relative abundances of different strains did differ between skin states (Fig. 3C) [62]. Additionally, genomic comparison of 71 different P. acnes strains shows that the acne-associated genomes maintained different chromosomal genomic region loci and a linear plasmid, thereby suggesting that there may be specific genes at these loci that contribute to acne pathology [62].

These findings suggest that strain-level analysis of the skin microbiota may be instrumental in explaining disease pathogenesis.

2.5.4 Dandruff

Dandruff is a mild inflammatory condition that is characterized by scaling of skin on the scalp. Malassezia fungi were proposed as the primary cause of dandruff in 1874, and this idea is still prevalent today. In fact, dandruff therapeutic shampoos are made with strong antifungal compounds in an attempt to target fungal causes of the disease [63]. Although Malassezia is the dominant fungal genus cultured from the skin and has been shown to increase in abundance on dandruff-afflicted skin [64], recent work has suggested that the dandruff microbial communities are more complex. A molecular analysis of the 26S rRNA gene of the fungal communities associated with healthy and dandruff-afflicted scalps showed that, similar to what was shown in previous studies, the relative abundance of Malassezia was increased in the dandruff-afflicted scalp skin (Fig. 3D) [23]. The study also reported that dandruff-afflicted skin harbored increased relative abundances of Penicillium and Filoblasidium floriforme that correlated with increased severity of dandruff. Furthermore, because Malassezia is found in the commensal fungal microbiota, it is not a likely cause of disease on its own, and thus there may be other interactive mechanisms involved in etiology.

2.5.5 Merkel Cell Carcinoma

Merkel cell carcinoma (MCC) is a rare but aggressive malignant, neuroendocrine tumor that has been increasing in incidence in past decades [65]. In 2008, Feng et al. (2008) showed that there was a novel virus (Merkel cell polyomavirus [MCPyV]) associated with MCC tissue, but not healthy tissue. Numerous additional studies, across diverse populations, also found strong associations between the presence of MCPyV and MCC [66-75]. Investigation into the virus's

role in health and disease have shown that MCPyVs are members of the commensal skin viral communities, are often asymptomatically carried, and can be shed from the skin as assembled virus particles [66-76]. Because this virus is a commensal microbe on healthy skin that does not develop MCC, there are likely other factors that interact with MCPyV to cause disease, including host immune function.

2.6 Diagnostic and therapeutic potential of the microbiome in skin disease

It is clear that the microbiome plays a broad, intricate, and complicated role in both human skin health and disease. In light of the many translational opportunities to use these findings in the clinic, a great amount of research has been devoted to clinical applications of microbiome research (Table 1). Probiotics, live microorganisms or microorganism components that confer health benefits, have long been administered therapeutically and prophylactically to the gastrointestinal tract, even before their mechanism was known. Effective and safe probiotics for use on the skin is an area of active investigation with great promise [77-79]. For those skin diseases that may be influenced by the gut microbiota, there is evidence that probiotic intervention may provide benefit. The efficacy of probiotics in treating AD remains somewhat controversial, but evidence suggests that administration of some Lactobacillus rhamnosus strains to mothers before and after birth reduces the occurrence and frequency of infantile AD [80-83].

Another microbiome-related approach to treating disease may be the use of prebiotics, which consist of substrates that promote the growth and/or metabolic activity of beneficial indigenous microbiota. Current prebiotics are primarily associated with ingestion and consequent manipulation of the gut microbiome. Different types of gut prebiotics such as galacto- and long-chain fructo-oligosaccharides show promise in treating infants with AD [84, 85]. But imaginable prebiotic approaches such as treating the skin with substrates to alter the environmental

conditions and thus promote or discourage the growth of certain microbiota may offer promise for the treatment of skin disorders whose pathogenesis is clearly linked to a microbial cause.

Genetic engineering of microorganisms as vectors for delivery of therapeutic genes is another area of active investigation. The potential utility of such approaches was shown by a study in which E. coli was engineered to express a quorum-sensing peptide that is naturally expressed by Vibrio cholera and inhibits V. cholera virulence [86]. Administration of the genetically engineered microbe to the gut of infant mice before challenge with V. cholera significantly increased survival while decreasing cholera toxin binding to the intestines. Bacteriophage can also be engineered and administered for therapeutic benefit. For example, bacteriophage have been used to deliver gene constructs to reverse antibiotic resistance in bacteria populations [87]. This approach provides the first steps toward applying evolutionary pressure against antibiotic resistance, while reversing the pressure toward antibiotic resistance from decades of antibiotic use. Of particular concern to the skin is multidrug resistance in skin-associated opportunistic pathogens, such as S. epidermidis, S. aureus, and Pseudomonas aeruginosa [88, 89]. Bacteriophages have also been engineered to degrade bacterial biofilms [90], a significant therapeutic challenge because antibiotics are often not able to physically access the bacteria comprising the biofilm, and are therefore not effective in treating them.

An in-depth understanding of the human skin microbiota may also have important implications in informing synthetic biology therapeutics. For example, comparative genomic analysis of P. acnes bacteriophage genomes led to the discovery of a highly conserved gene-encoding endolysin, an enzyme with broad lytic potential for P. acnes hosts [36]. The utility of endolysin as an antibacterial has been shown in other phage–host systems, and bacterial resistance to the recombinant protein was not observed even after repeated exposure [91].

2.7 Concluding remarks

The skin acts as both a protective physical barrier between the body and the external environment, as well as an environmental substrate that harbors rich and diverse communities of microorganisms that contribute to skin health and disease. The recent advent of molecular and metagenomic techniques for microbial community analysis has addressed many culture-based limitations. As a result, a greater appreciation of the microbial diversity across different skin sites as well as the diversity between people, over time, has emerged. It is becoming increasingly apparent that certain microbes promote healthy skin equilibrium, and contemporary molecular approaches have also provided greater information about the role of microbial community disturbances in disease pathogenesis.

The commensal fungal and viral communities in either health or disease remain largely uncharacterized, and future investigations are likely to focus on these knowledge gaps. Most molecular studies up to this point have focused on taxonomic characterization of microbial skin communities. Although this approach is valuable because taxonomy provides a functional proxy for patterns of the genomes present, new techniques will allow for more in-depth investigations, beyond taxonomic identification. In light of increasingly feasible whole metagenomic shotgun sequencing approaches, investigations will be able to focus directly on the genetic functional potential and assess the community compositions of relevant genes. We also expect that, with the ever-advancing technologies and bioinformatics associated with mRNA sequencing (the metatranscriptome) and protein community analysis (the metaproteome), significant effort will be directed toward the functional aspects of microbiomes associated with skin health and disease.

Finally, a looming challenge is applying this knowledge to develop therapeutic and diagnostic tools for the clinic. Enhanced understanding of the skin microbiome will continue to inform research toward probiotic and prebiotic development, prevention of antibiotic resistance gene transfer, bacteriophage-mediated treatments, and gene delivery using bacterial vectors. New therapeutic developments will allow for a type of "microbiome engineering" in which the community composition will be stimulated and/or manipulated to include beneficial components. Additionally, in light of increasing antibiotic resistance across medically relevant bacterial populations, there will likely be an increased interest in alternative approaches to treating infections, as well as slowing the spread of resistance.

2.8 Acknowledgements

We thank members of the Grice laboratory for their underlying contributions. G.D.H. is supported by the Department of Defense through the National Defense Science and Engineering Graduate Fellowship Program. E.A.G. gratefully acknowledges the support of the National Institutes of Health (AR060873) and the University of Pennsylvania Skin Disease Research Center (supported by AR057217).

2.9 Figures

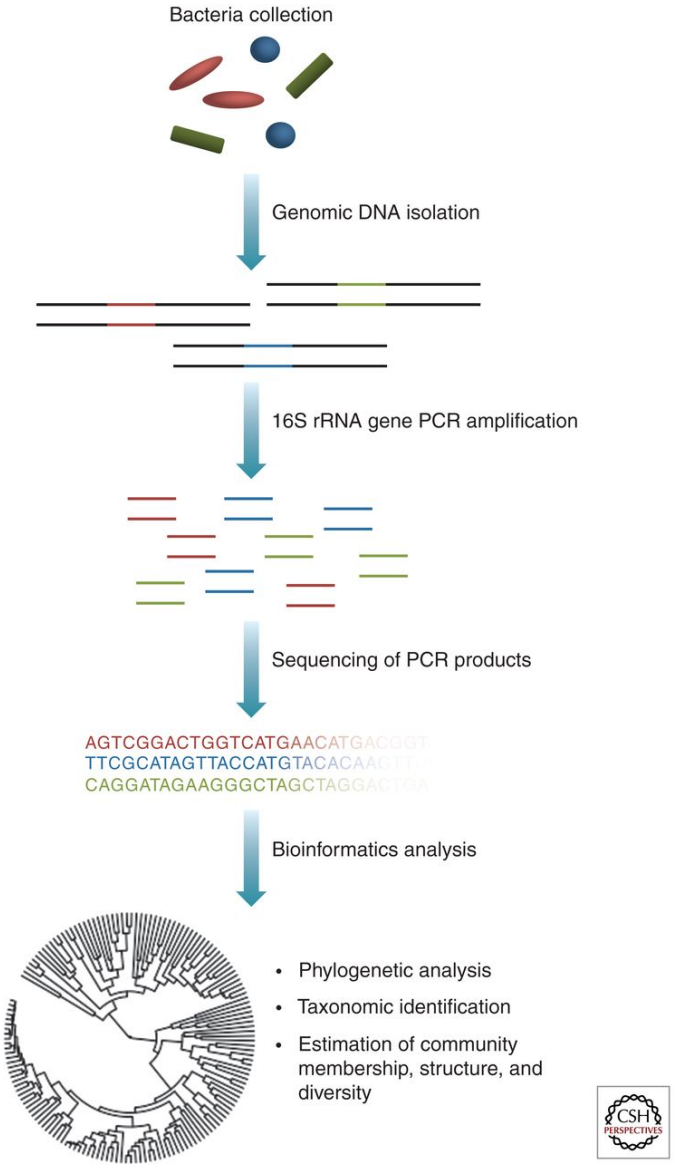

Figure 1

The workflow of a bacterial 16S rRNA gene microbiome study. A heterogeneous mixture of genomic DNA is extracted from samples taken from the skin. Primers, containing barcodes that allow for multiplexing, are designed to the desired region of the 16S rRNA gene. 16S rRNA gene PCR products are amplified and sequenced. Low-quality sequences are removed, and various

analyses are performed. These analyses can include assignment to taxonomy, analysis of shared phylogeny, and analysis of microbial community membership, structure, and diversity.

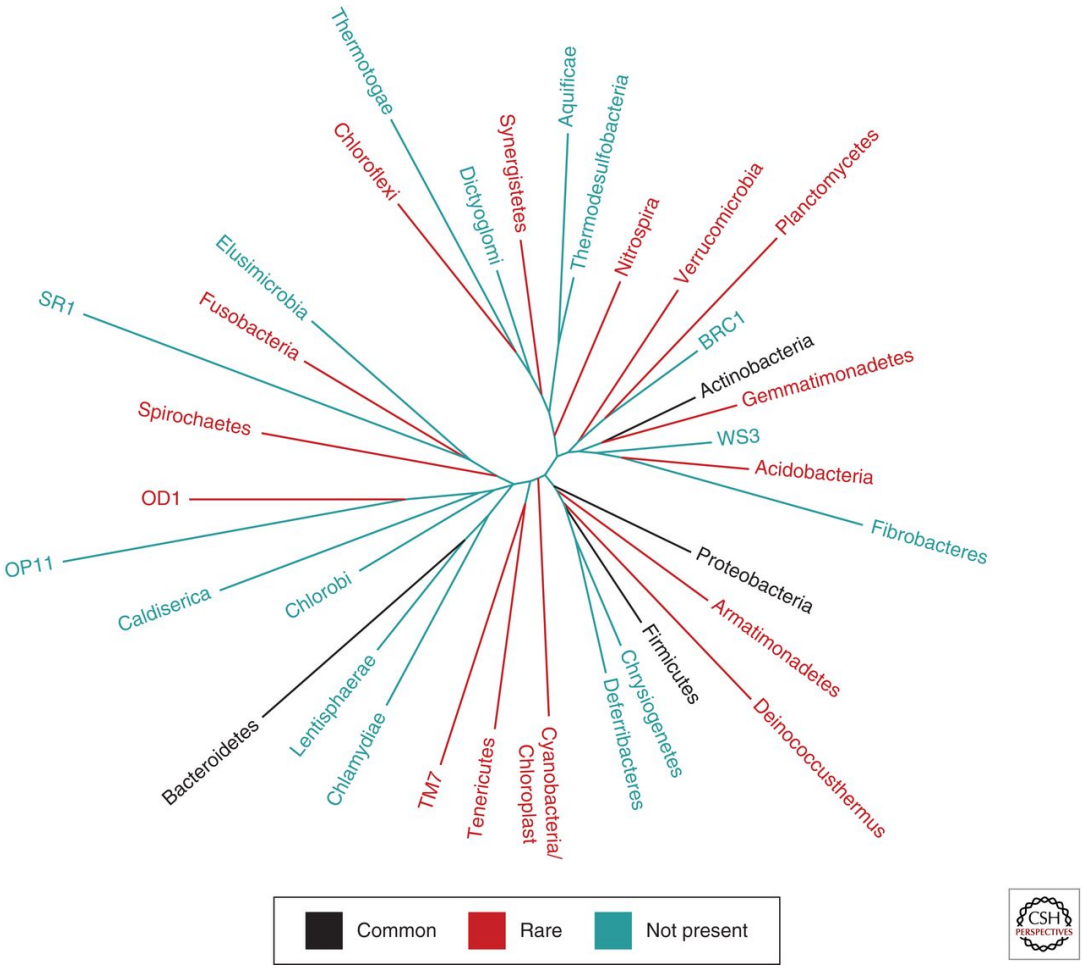

Figure 2

Bacterial diversity of the skin. Phylogenetic tree of the domain Bacteria with each branch representing a phylum. Black branches represent numerically abundant phyla on the skin, red branches represent rare phyla on the skin, and green branches represent phyla that are absent from the skin. (The data are derived from Grice et al. 2009.)

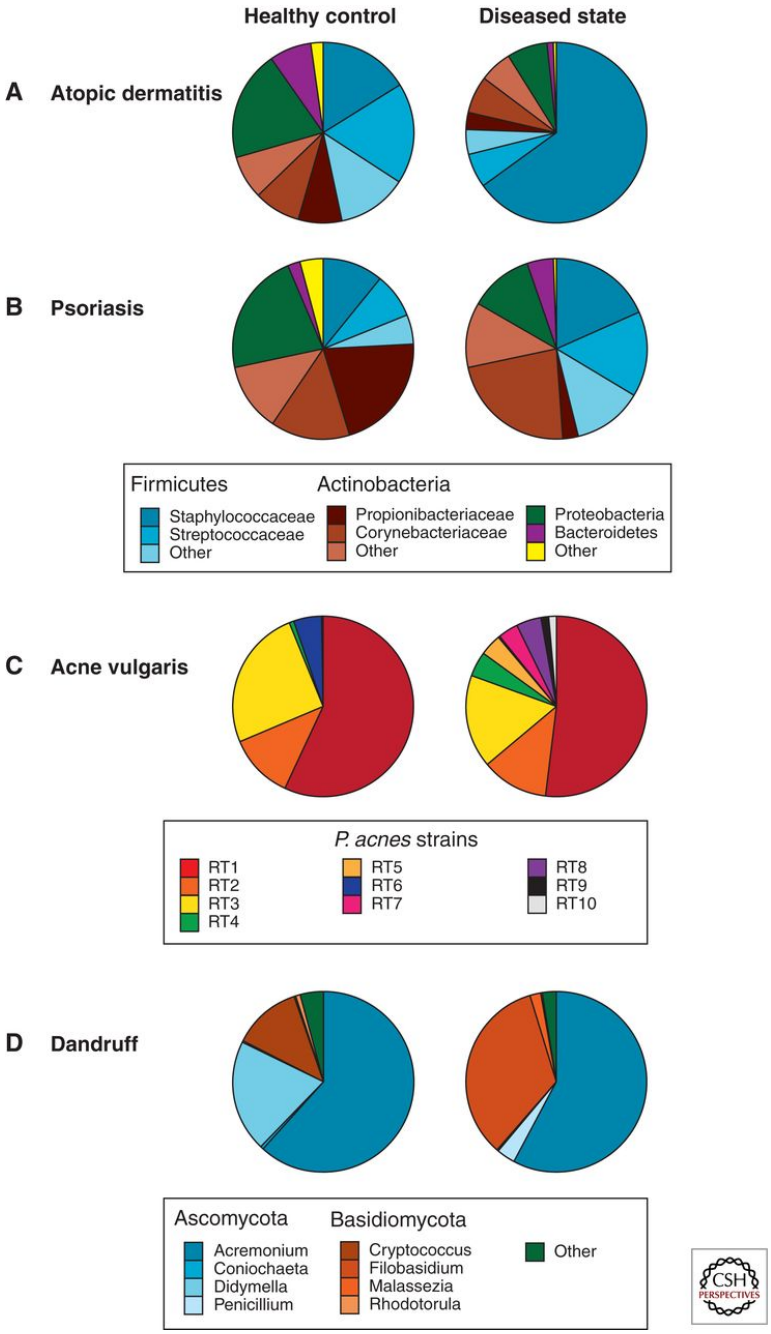

Figure 3

Changes in skin microbiota are associated with disease. (A) Relative abundance of bacteria (16S rRNA) in 12 children with AD flares as compared with 11 healthy controls (Kong et al. 2012). (B) Relative abundance of bacteria (16S rRNA) in six patients with psoriasis, in the lesional area as compared with unaffected skin as a control (Gao et al. 2008). (C) Relative abundance of P.

acnes strains in 49 acne patients and 52 healthy individuals (Fitz-Gibbon et al. 2013). (D) Relative abundance of fungi (26S rRNA) in three healthy scalps and four dandruff-afflicted scalps (Park et al. 2012).

2.10 Tables

Therapeutic	Disease target example	Summary
Probiotics	Atopic dermatitis (Batchelor et al. 2010)	Administration of microorganisms, or their components, to confer health benefits
Prebiotics	Atopic dermatitis (Foolad et al. 2008)	Administration of a substance to promote growth and/or action of therapeutically beneficial indigenous microbes
Bacteria-mediated gene delivery	Vibrio cholera infection (Duan and March 2010)	Inhibition of pathogen virulence by administrating bacteria engineered to express genes absent from the commensal community
Bacteriophage-mediated antibiotic susceptibility	Antibiotic-resistant bacteria (Edgar et al. 2012)	Reduction in antibiotic-resistant bacteria populations by using bacteriophage to deliver gene constructs that promote evolution toward antibiotic susceptibility
Bacteriophage antimicrobial peptides	*P. acnes* infection (Marinelli et al. 2012)	Use of bacteriophage peptides, such as endolysin, which promotes bacteria lysis during the lytic bacteriophage life cycle, against bacterial infections
Direct bacteriophage therapy	*Escherichia coli* biofilm infections (Lu and Collins 2007)	Administration of bacteriophages, which can be engineered to express enzymes for biofilm destruction, to combat biofilm-forming bacteria infections

Table 1

Therapeutic approaches based on the microbiome.

2.11 References

1. Grice EA, Segre JA: **The human microbiome: our second genome.** *Annu Rev Genomics Hum Genet* 2012, **13:**151-170.
2. Fuchs E, Raghavan S: **Getting under the skin of epidermal morphogenesis.** *Nat Rev Genet* 2002, **3:**199-209.
3. Nakatsuji T, Gallo RL: **Antimicrobial peptides: old molecules with new ideas.** *J Invest Dermatol* 2012, **132:**887-895.
4. Lai Y, Di Nardo A, Nakatsuji T, Leichtle A, Yang Y, Cogen AL, Wu ZR, Hooper LV, Schmidt RR, von Aulock S, et al: **Commensal bacteria regulate Toll-like receptor 3-dependent inflammation after skin injury.** *Nat Med* 2009, **15:**1377-1382.
5. Chehoud C, Rafail S, Tyldsley AS, Seykora JT, Lambris JD, Grice EA: **Complement modulates the cutaneous microbiome and inflammatory milieu.** *Proc Natl Acad Sci U S A* 2013, **110:**15061-15066.
6. Naik S, Bouladoux N, Wilhelm C, Molloy MJ, Salcedo R, Kastenmuller W, Deming C, Quinones M, Koo L, Conlan S, et al: **Compartmentalized control of skin immunity by resident commensals.** *Science* 2012, **337:**1115-1119.
7. MJ M: *The ecology of the human skin.* Thomas, Springfield, IL.; 1965.
8. Lane DJ, Pace B, Olsen GJ, Stahl DA, Sogin ML, Pace NR: **Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses.** *Proc Natl Acad Sci U S A* 1985, **82:**6955-6959.
9. Chase MW, Fay MF: **Ecology. Barcoding of plants and fungi.** *Science* 2009, **325:**682-683.
10. Metagenomics NRCCo: **The new science of metagenomics: Revealing the secrets of our microbial planet.** *The National Academies Press, Washington, DC* 2007.
11. Evans CA, Smith WM, Johnston EA, Giblett ER: **Bacterial flora of the normal human skin.** *J Invest Dermatol* 1950, **15:**305-324.
12. Grice EA, Kong HH, Conlan S, Deming CB, Davis J, Young AC, Bouffard GG, Blakesley RW, Murray PR, Green ED, et al: **Topographical and temporal diversity of the human skin microbiome.** *Science* 2009, **324:**1190-1192.
13. Gao Z, Tseng CH, Pei Z, Blaser MJ: **Molecular analysis of human forearm superficial skin bacterial biota.** *Proc Natl Acad Sci U S A* 2007, **104:**2927-2932.
14. Grice EA, Kong HH, Renaud G, Young AC, Bouffard GG, Blakesley RW, Wolfsberg TG, Turner ML, Segre JA: **A diversity profile of the human skin microbiota.** *Genome Res* 2008, **18:**1043-1050.
15. Blaser MJ, Dominguez-Bello MG, Contreras M, Magris M, Hidalgo G, Estrada I, Gao Z, Clemente JC, Costello EK, Knight R: **Distinct cutaneous bacterial assemblages in a sampling of South American Amerindians and US residents.** *ISME J* 2013, **7:**85-95.
16. Fierer N, Hamady M, Lauber CL, Knight R: **The influence of sex, handedness, and washing on the diversity of hand surface bacteria.** *Proc Natl Acad Sci U S A* 2008, **105:**17994-17999.
17. Staudinger T, Pipal A, Redl B: **Molecular analysis of the prevalent microbiota of human male and female forehead skin compared to forearm skin and the influence of make-up.** *J Appl Microbiol* 2011, **110:**1381-1389.
18. Capone KA, Dowd SE, Stamatas GN, Nikolovski J: **Diversity of the human skin microbiome early in life.** *J Invest Dermatol* 2011, **131:**2026-2032.
19. Oh J, Conlan S, Polley EC, Segre JA, Kong HH: **Shifts in human skin and nares microbiota of healthy children and adults.** *Genome Med* 2012, **4:**77.

20. Consortium HMP: **Structure, function and diversity of the healthy human microbiome.** *Nature* 2012, **486:**207-214.
21. Gaitanis G, Magiatis P, Hantschke M, Bassukas ID, Velegraki A: **The Malassezia genus in skin and systemic diseases.** *Clin Microbiol Rev* 2012, **25:**106-141.
22. Paulino LC, Tseng CH, Blaser MJ: **Analysis of Malassezia microbiota in healthy superficial human skin and in psoriatic lesions by multiplex real-time PCR.** *FEMS Yeast Res* 2008, **8:**460-471.
23. Park HK, Ha MH, Park SG, Kim MN, Kim BJ, Kim W: **Characterization of the fungal microbiota (mycobiome) in healthy and dandruff-afflicted human scalps.** *PLoS One* 2012, **7:**e32847.
24. Findley K, Oh J, Yang J, Conlan S, Deming C, Meyer JA, Schoenfeld D, Nomicos E, Park M, Program NIHISCCS, et al: **Topographic diversity of fungal and bacterial communities in human skin.** *Nature* 2013, **498:**367-370.
25. de Villiers EM, Lavergne D, McLaren K, Benton EC: **Prevailing papillomavirus types in non-melanoma carcinomas of the skin in renal allograft recipients.** *Int J Cancer* 1997, **73:**356-361.
26. Astori G, Lavergne D, Benton C, Hockmayr B, Egawa K, Garbe C, de Villiers EM: **Human papillomaviruses are commonly found in normal skin of immunocompetent hosts.** *J Invest Dermatol* 1998, **110:**752-755.
27. Antonsson A, Forslund O, Ekberg H, Sterner G, Hansson BG: **The ubiquity and impressive genomic diversity of human skin papillomaviruses suggest a commensalic nature of these viruses.** *J Virol* 2000, **74:**11636-11641.
28. Antonsson A, Erfurt C, Hazard K, Holmgren V, Simon M, Kataoka A, Hossain S, Hakangard C, Hansson BG: **Prevalence and type spectrum of human papillomaviruses in healthy skin samples collected in three continents.** *J Gen Virol* 2003, **84:**1881-1886.
29. Antonsson A, Karanfilovska S, Lindqvist PG, Hansson BG: **General acquisition of human papillomavirus infections of skin occurs in early infancy.** *J Clin Microbiol* 2003, **41:**2509-2514.
30. Forslund O: **Genetic diversity of cutaneous human papillomaviruses.** *J Gen Virol* 2007, **88:**2662-2669.
31. Moens U, Ludvigsen M, Van Ghelue M: **Human polyomaviruses in skin diseases.** *Patholog Res Int* 2011, **2011:**123491.
32. Schowalter RM, Pastrana DV, Pumphrey KA, Moyer AL, Buck CB: **Merkel cell polyomavirus and two previously unknown polyomaviruses are chronically shed from human skin.** *Cell Host Microbe* 2010, **7:**509-515.
33. Foulongne V, Sauvage V, Hebert C, Dereure O, Cheval J, Gouilh MA, Pariente K, Segondy M, Burguiere A, Manuguerra JC, et al: **Human skin microbiota: high diversity of DNA viruses identified on the human skin by high throughput sequencing.** *PLoS One* 2012, **7:**e38499.
34. Kwan T, Liu J, DuBow M, Gros P, Pelletier J: **The complete genomes and proteomes of 27 Staphylococcus aureus bacteriophages.** *Proc Natl Acad Sci U S A* 2005, **102:**5174-5179.
35. Goerke C, Pantucek R, Holtfreter S, Schulte B, Zink M, Grumann D, Broker BM, Doskar J, Wolz C: **Diversity of prophages in dominant Staphylococcus aureus clonal lineages.** *J Bacteriol* 2009, **191:**3462-3468.
36. Marinelli LJ, Fitz-Gibbon S, Hayes C, Bowman C, Inkeles M, Loncaric A, Russell DA, Jacobs-Sera D, Cokus S, Pellegrini M, et al: **Propionibacterium acnes**

bacteriophages display limited genetic diversity and broad killing activity against bacterial skin isolates. *MBio* 2012, **3**.
37. Wang X, Kim Y, Ma Q, Hong SH, Pokusaeva K, Sturino JM, Wood TK: **Cryptic prophages help bacteria cope with adverse environments.** *Nat Commun* 2010, **1:**147.
38. Fancello L, Desnues C, Raoult D, Rolain JM: **Bacteriophages and diffusion of genes encoding antimicrobial resistance in cystic fibrosis sputum microbiota.** *J Antimicrob Chemother* 2011, **66:**2448-2454.
39. Minot S, Sinha R, Chen J, Li H, Keilbaugh SA, Wu GD, Lewis JD, Bushman FD: **The human gut virome: inter-individual variation and dynamic response to diet.** *Genome Res* 2011, **21:**1616-1625.
40. Nakaminami H, Noguchi N, Nishijima S, Kurokawa I, So H, Sasatsu M: **Transduction of the plasmid encoding antiseptic resistance gene qacB in Staphylococcus aureus.** *Biol Pharm Bull* 2007, **30:**1412-1415.
41. Varga M, Kuntova L, Pantucek R, Maslanova I, Ruzickova V, Doskar J: **Efficient transfer of antibiotic resistance plasmids by transduction within methicillin-resistant Staphylococcus aureus USA300 clone.** *FEMS Microbiol Lett* 2012, **332:**146-152.
42. Huang JT, Abrams M, Tlougan B, Rademaker A, Paller AS: **Treatment of Staphylococcus aureus colonization in atopic dermatitis decreases disease severity.** *Pediatrics* 2009, **123:**e808-814.
43. Yao Y, Kozman A, Al-Hassani M, Saha CK, Yi Q, Yao W, Mousdicas N, Kaplan MH, Travers JB: **Identification of staphylococcal protein A in infected atopic dermatitis lesions.** *J Invest Dermatol* 2010, **130:**2502-2504.
44. Travers JB, Kozman A, Mousdicas N, Saha C, Landis M, Al-Hassani M, Yao W, Yao Y, Hyatt AM, Sheehan MP, et al: **Infected atopic dermatitis lesions contain pharmacologic amounts of lipoteichoic acid.** *J Allergy Clin Immunol* 2010, **125:**146-152 e141-142.
45. Kong HH, Oh J, Deming C, Conlan S, Grice EA, Beatson MA, Nomicos E, Polley EC, Komarow HD, Murray PR, et al: **Temporal shifts in the skin microbiome associated with disease flares and treatment in children with atopic dermatitis.** *Genome Res* 2012, **22:**850-859.
46. Kaga M, Sugita T, Nishikawa A, Wada Y, Hiruma M, Ikeda S: **Molecular analysis of the cutaneous Malassezia microbiota from the skin of patients with atopic dermatitis of different severities.** *Mycoses* 2011, **54:**e24-28.
47. Zhang E, Tanaka T, Tajima M, Tsuboi R, Nishikawa A, Sugita T: **Characterization of the skin fungal microbiota in patients with atopic dermatitis and in healthy subjects.** *Microbiol Immunol* 2011, **55:**625-632.
48. Wang M, Karlsson C, Olsson C, Adlerberth I, Wold AE, Strachan DP, Martricardi PM, Aberg N, Perkin MR, Tripodi S, et al: **Reduced diversity in the early fecal microbiota of infants with atopic eczema.** *J Allergy Clin Immunol* 2008, **121:**129-134.
49. Penders J, Stobberingh EE, Thijs C, Adams H, Vink C, van Ree R, van den Brandt PA: **Molecular fingerprinting of the intestinal microbiota of infants in whom atopic eczema was or was not developing.** *Clin Exp Allergy* 2006, **36:**1602-1608.
50. Penders J, Thijs C, van den Brandt PA, Kummeling I, Snijders B, Stelma F, Adams H, van Ree R, Stobberingh EE: **Gut microbiota composition and development of atopic manifestations in infancy: the KOALA Birth Cohort Study.** *Gut* 2007, **56:**661-667.

51. Sandilands A, Terron-Kwiatkowski A, Hull PR, O'Regan GM, Clayton TH, Watson RM, Carrick T, Evans AT, Liao H, Zhao Y, et al: **Comprehensive analysis of the gene encoding filaggrin uncovers prevalent and rare mutations in ichthyosis vulgaris and atopic eczema.** *Nat Genet* 2007, **39:**650-654.
52. Scharschmidt TC, List K, Grice EA, Szabo R, Program NCS, Renaud G, Lee CC, Wolfsberg TG, Bugge TH, Segre JA: **Matriptase-deficient mice exhibit ichthyotic skin with a selective shift in skin microbiota.** *J Invest Dermatol* 2009, **129:**2435-2442.
53. Gottlieb AB: **Therapeutic options in the treatment of psoriasis and atopic dermatitis.** *J Am Acad Dermatol* 2005, **53:**S3-16.
54. Boehncke WH, Dressel D, Zollner TM, Kaufmann R: **Pulling the trigger on psoriasis.** *Nature* 1996, **379:**777.
55. Aly R, Maibach HE, Mandel A: **Bacterial flora in psoriasis.** *Br J Dermatol* 1976, **95:**603-606.
56. Rosenberg EW, Noah PW, Skinner RB, Jr.: **Microorganisms and psoriasis.** *J Natl Med Assoc* 1994, **86:**305-310.
57. Nickoloff BJ, Schroder JM, von den Driesch P, Raychaudhuri SP, Farber EM, Boehncke WH, Morhenn VB, Rosenberg EW, Schon MP, Holick MF: **Is psoriasis a T-cell disease?** *Exp Dermatol* 2000, **9:**359-375.
58. Paulino LC, Tseng CH, Strober BE, Blaser MJ: **Molecular analysis of fungal microbiota in samples from healthy human skin and psoriatic lesions.** *J Clin Microbiol* 2006, **44:**2933-2941.
59. Gao Z, Tseng CH, Strober BE, Pei Z, Blaser MJ: **Substantial alterations of the cutaneous bacterial biota in psoriatic lesions.** *PLoS One* 2008, **3:**e2719.
60. Fahlen A, Engstrand L, Baker BS, Powles A, Fry L: **Comparison of bacterial microbiota in skin biopsies from normal and psoriatic skin.** *Arch Dermatol Res* 2012, **304:**15-22.
61. Tan HH: **Antibacterial therapy for acne: a guide to selection and use of systemic agents.** *Am J Clin Dermatol* 2003, **4:**307-314.
62. Fitz-Gibbon S, Tomida S, Chiu BH, Nguyen L, Du C, Liu M, Elashoff D, Erfe MC, Loncaric A, Kim J, et al: **Propionibacterium acnes strain populations in the human skin microbiome associated with acne.** *J Invest Dermatol* 2013, **133:**2152-2160.
63. Bulmer AC, Bulmer GS: **The antifungal action of dandruff shampoos.** *Mycopathologia* 1999, **147:**63-65.
64. McGinley KJ, Leyden JJ, Marples RR, Kligman AM: **Quantitative microbiology of the scalp in non-dandruff, dandruff, and seborrheic dermatitis.** *J Invest Dermatol* 1975, **64:**401-405.
65. Hodgson NC: **Merkel cell carcinoma: changing incidence trends.** *J Surg Oncol* 2005, **89:**1-4.
66. Foulongne V, Kluger N, Dereure O, Brieu N, Guillot B, Segondy M: **Merkel cell polyomavirus and Merkel cell carcinoma, France.** *Emerg Infect Dis* 2008, **14:**1491-1493.
67. Kassem A, Schopflin A, Diaz C, Weyers W, Stickeler E, Werner M, Zur Hausen A: **Frequent detection of Merkel cell polyomavirus in human Merkel cell carcinomas and identification of a unique deletion in the VP1 gene.** *Cancer Res* 2008, **68:**5009-5013.

68. Becker JC, Houben R, Ugurel S, Trefzer U, Pfohler C, Schrama D: **MC polyomavirus is frequently present in Merkel cell carcinoma of European patients.** *J Invest Dermatol* 2009, **129:**248-250.
69. Duncavage EJ, Zehnbauer BA, Pfeifer JD: **Prevalence of Merkel cell polyomavirus in Merkel cell carcinoma.** *Mod Pathol* 2009, **22:**516-521.
70. Katano H, Ito H, Suzuki Y, Nakamura T, Sato Y, Tsuji T, Matsuo K, Nakagawa H, Sata T: **Detection of Merkel cell polyomavirus in Merkel cell carcinoma and Kaposi's sarcoma.** *J Med Virol* 2009, **81:**1951-1958.
71. Touze A, Gaitan J, Maruani A, Le Bidre E, Doussinaud A, Clavel C, Durlach A, Aubin F, Guyetant S, Lorette G, Coursaget P: **Merkel cell polyomavirus strains in patients with merkel cell carcinoma.** *Emerg Infect Dis* 2009, **15:**960-962.
72. Mangana J, Dziunycz P, Kerl K, Dummer R, Cozzio A: **Prevalence of Merkel cell polyomavirus among Swiss Merkel cell carcinoma patients.** *Dermatology* 2010, **221:**184-188.
73. Jung HS, Choi YL, Choi JS, Roh JH, Pyon JK, Woo KJ, Lee EH, Jang KT, Han J, Park CS, et al: **Detection of Merkel cell polyomavirus in Merkel cell carcinomas and small cell carcinomas by PCR and immunohistochemistry.** *Histol Histopathol* 2011, **26:**1231-1241.
74. Paolini F, Donati P, Amantea A, Bucher S, Migliano E, Venuti A: **Merkel cell polyomavirus in Merkel cell carcinoma of Italian patients.** *Virol J* 2011, **8:**103.
75. Haitz KA, Rady PL, Nguyen HP, He Q, Prieto VG, Tyring SK, Ciurea AM: **Merkel cell polyomavirus DNA detection in a patient with Merkel cell carcinoma and multiple other skin cancers.** *Int J Dermatol* 2012, **51:**442-444.
76. Foulongne V, Dereure O, Kluger N, Moles JP, Guillot B, Segondy M: **Merkel cell polyomavirus DNA detection in lesional and nonlesional skin from patients with Merkel cell carcinoma or other skin diseases.** *Br J Dermatol* 2010, **162:**59-63.
77. Muizzuddin N, Maher W, Sullivan M, Schnittger S, Mammone T: **Physiological effect of a probiotic on skin.** *J Cosmet Sci* 2012, **63:**385-395.
78. Lew LC, Liong MT: **Bioactives from probiotics for dermal health: functions and benefits.** *J Appl Microbiol* 2013.
79. Shu M, Wang Y, Yu J, Kuo S, Coda A, Jiang Y, Gallo RL, Huang CM: **Fermentation of Propionibacterium acnes, a commensal bacterium in the human skin microbiome, as skin probiotics against methicillin-resistant Staphylococcus aureus.** *PLoS One* 2013, **8:**e55380.
80. Kalliomaki M, Salminen S, Arvilommi H, Kero P, Koskinen P, Isolauri E: **Probiotics in primary prevention of atopic disease: a randomised placebo-controlled trial.** *Lancet* 2001, **357:**1076-1079.
81. Wickens K, Black PN, Stanley TV, Mitchell E, Fitzharris P, Tannock GW, Purdie G, Crane J: **A differential effect of 2 probiotics in the prevention of eczema and atopy: a double-blind, randomized, placebo-controlled trial.** *J Allergy Clin Immunol* 2008, **122:**788-794.
82. Wickens K, Black P, Stanley TV, Mitchell E, Barthow C, Fitzharris P, Purdie G, Crane J: **A protective effect of Lactobacillus rhamnosus HN001 against eczema in the first 2 years of life persists to age 4 years.** *Clin Exp Allergy* 2012, **42:**1071-1079.
83. Dotterud CK, Storro O, Johnsen R, Oien T: **Probiotics in pregnant women to prevent allergic disease: a randomized, double-blind trial.** *Br J Dermatol* 2010, **163:**616-623.

84. Moroi M, Uchi S, Nakamura K, Sato S, Shimizu N, Fujii M, Kumagai T, Saito M, Uchiyama K, Watanabe T, et al: **Beneficial effect of a diet containing heat-killed Lactobacillus paracasei K71 on adult type atopic dermatitis.** *J Dermatol* 2011, **38:**131-139.
85. Arslanoglu S, Moro GE, Schmitt J, Tandoi L, Rizzardi S, Boehm G: **Early dietary intervention with a mixture of prebiotic oligosaccharides reduces the incidence of allergic manifestations and infections during the first two years of life.** *J Nutr* 2008, **138:**1091-1095.
86. Duan F, March JC: **Engineered bacterial communication prevents Vibrio cholerae virulence in an infant mouse model.** *Proc Natl Acad Sci U S A* 2010, **107:**11260-11264.
87. Edgar R, Friedman N, Molshanski-Mor S, Qimron U: **Reversing bacterial resistance to antibiotics by phage-mediated delivery of dominant sensitive genes.** *Appl Environ Microbiol* 2012, **78:**744-751.
88. Branski LK, Al-Mousawi A, Rivero H, Jeschke MG, Sanford AP, Herndon DN: **Emerging infections in burns.** *Surg Infect (Larchmt)* 2009, **10:**389-397.
89. Otto M: **Staphylococcus epidermidis--the 'accidental' pathogen.** *Nat Rev Microbiol* 2009, **7:**555-567.
90. Lu TK, Collins JJ: **Dispersing biofilms with engineered enzymatic bacteriophage.** *Proc Natl Acad Sci U S A* 2007, **104:**11197-11202.
91. Fischetti VA: **Bacteriophage lysins as effective antibacterials.** *Curr Opin Microbiol* 2008, **11:**393-400.

CHAPTER 3 – The Human Skin dsDNA Virome: Topographical and Temporal Diversity, Genetic Enrichment, and Dynamic Associations with the Host Microbiome

The contents of this chapter are under review for publication as:

Hannigan GD, Meisel JS, Tyldsley AS, Zheng Q, Hodkinson BP, SanMiguel AJ, Minot S, Bushman FD, Grice EA. (In Review)

3.1 Abstract

Viruses comprise a major component of the human microbiota, but are poorly understood in the skin, our primary barrier to the external environment. Viral communities have the potential to modulate states of cutaneous health and disease. Bacteriophages are known to influence the structure and function of microbial communities through predation and genetic exchange. Human viruses are associated with skin cancers and a multitude of cutaneous manifestations. Despite these important roles, little is known regarding the human skin virome and its interactions with the host microbiome. Here we evaluate the human cutaneous dsDNA virome by metagenomic sequencing of DNA from purified virus-like particles (VLPs). In parallel we employ metagenomic sequencing of the total skin microbiome to assess co-variation and infer interactions with the virome. Samples were collected from sixteen subjects at eight body sites over one month. In addition to microenviroment, which is known to partition bacterial and fungal microbiota, natural skin occlusion was strongly associated with skin virome community composition. Viral contigs were enriched for genes indicative of a temperate phage replication style, and also maintained genes encoding potential antibiotic resistance and virulence factors. CRISPR spacers identified in the bacterial DNA sequences provided a record of phage predation and suggest a mechanism to explain spatial partitioning of skin phage communities. Finally, we model the structure of bacterial and phage communities together to reveal a complex microbial

environment with a *Corynebacterium* hub. These results reveal the previously underappreciated diversity, encoded functions, and viral-microbial dynamic unique to the human skin virome.

3.2 Importance

To date, most cutaneous microbiome studies have focused on bacterial and fungal communities. Skin viral communities and their relationships with their hosts remain poorly understood despite their potential to modulate states of cutaneous health and disease. Previous studies employing whole metagenome sequencing without purification for virus-like particles (VLPs) have provided some insight into the viral component of the skin microbiome, but have not completely characterized these communities or analyzed interactions with the host microbiome. Additionally, these studies had potentially biased community representations, shallower coverage of the viral community components, and an inability to assess the viral dark matter of the skin. Here we present the first optimized virus purification technique and analysis tools for gaining novel insights into the skin virome and its potential interactions with the host microbiome.

3.3 Introduction

The human skin is a barrier to the external environment and home to diverse and distinctive microbial communities. To date, most cutaneous microbiome studies have focused on bacterial and fungal communities, their modulation of cutaneous immune responses, and the association of these microorganisms with dermatological disorders [1]. Recent metagenomic studies confirm the role of skin microenvironment and interpersonal variation in shaping the microbiome [2]. Skin viral communities and their relationships with their hosts remain poorly understood despite their potential to modulate states of cutaneous health and disease. Bacteriophages ("phages"; viruses that infect bacteria) can affect human health by altering the composition of their host bacterial communities through predation [3, 4]. Evidence of such dynamism is provided by acquisition and

diversification of bacterial clustered regularly interspaced short palindromic repeat (CRISPR) elements (e.g. [5]), which target phage genomes for destruction using nucleases guided by sequences encoded in the CRISPR arrays. Phages may also have long-term impacts on their hosts via lysogeny, in which phages integrate their genome into the host and adopt a quiescent state. New genes encoded on lysogens can affect host metabolism, virulence, antibiotic resistance, and sensitivity to other phages [6-9]. Phages may also serve as a genetic reservoir for bacterial adaptations during stress (i.e. antibiotic treatment) [10]. Viruses that replicate on human cells are also present in the skin and can affect human health, including Human Papillomaviruses (HPV), Human Polyomaviruses (HPyV), and Human Herpesviruses (HHV), and can cause skin cancers and other dermatological disorders.

Previous studies employing whole metagenome sequencing without purification for virus-like particles (VLPs) have provided some insight into the viral component of the skin microbiome, but have not completely characterized these communities or analyzed interactions with the host microbiome [2, 11, 12]. The study we present here employs gold standard techniques for the purification of viral DNA, thereby reducing contamination from human and bacterial cells, whose genomes are orders of magnitude longer than viral genomes. This allows for deeper viral sequencing and the use of reference-independent analyses to capture the impact of unknown or uncharacterized genomes, known as viral dark matter. We applied shotgun metagenomic analysis to purified VLPs, as well as unpurified whole skin microbial communities, conducting the first longitudinal, integrated analysis of the healthy human skin virome and the whole metagenome across diverse anatomical locations. The major questions we address with this novel dataset are: What is the biogeography and diversity of the human skin virome compared to the whole metagenome over time and across individuals? What genetic functions are encoded by the skin virome, including antibiotic resistance, virulence factors and auxiliary metabolic genes (AMGs;

"host" genes within phage genomes [13])? What can we infer about interactions between phages and their bacterial hosts, including the role of CRISPRs in maintaining virome community structure?

3.4 Results

3.4.1 Sampling, sequencing, and quality control

Cutaneous skin swabs were collected from 16 healthy volunteers with no known skin conditions between the ages of 23 and 53 years old (**Fig 1A-B**). Anatomical skin sites were sampled bilaterally (virome sample was collected at the site contralateral to the whole metagenome sample) and consisted of multiple diverse microenvironments: sebaceous (retroauricular crease [Ra], occiput [Oc], and forehead [Fh]), moist (axilla [Ax], toe web [Tw], and umbilicus [Um]), and intermittently moist (antecubital fossa [Ac] and palm [Pa]) (**Fig 1A**). Swab samples were collected at two time points separated by four weeks to assess stability of the communities.

After swabbing each subject's skin, we used one sample of the contralateral pair to purify and extract the VLP DNA using a protocol established for human and environmental viromes [14-16]. We extracted the DNA from the contralateral sample to investigate the whole microbial community, including bacterial, fungal, and viral members. Samples were prepared for shotgun sequencing on the Illumina MiSeq and HiSeq2500 platforms using the Illumina NexteraXT library preparation kit, which is designed for dsDNA. Therefore, our analysis focuses on dsDNA viruses and replicative intermediates of ssDNA viruses. Sample collection, sequence processing, and bioinformatics analyses are outlined in **Fig 1C**.

After quality filtering, the dsDNA virus dataset contained 260,714,906 total high quality sequence reads, with a median of 650,506 sequence reads per sample. The whole metagenome

dataset contained 368,341,329 total high quality sequence reads, with a median of 981,031 sequence reads per sample (See **Supplemental Fig 1A-D** and **Table 1** for sequence count statistics). Consistent with previous reports of similar human VLP preparations [15-18], a relaxed search against the entire NCBI non-redundant database revealed that 94.8% of VLP reads did not significantly match a known genome (blastn; E-value<10^{-3}), highlighting the importance of investigating viral dark matter. Similar classification identified 42.6% of the whole metagenome reads as unknown. In this study, we use multiple reference-independent approaches to address this subset of unclassified dark matter. The viral and whole metagenome datasets were independently assembled into contigs, and contigs >500 bp in length were selected for further analysis (See **Supplemental Fig 1E-H and Table 2** for contig coverage, count, and length statistics). Of these phage contigs, 9.0% were taxonomically identifiable, highlighting the utility of using contigs in taxonomy instead of using unaligned reads.

During each sampling event, we collected a blank negative control that never came into contact with skin. DNA was extracted from the control and sequenced in parallel with the experimental samples. Using the Bray-Curtis dissimilarity metric, we found significant separation of the control samples from the skin samples (**Supplemental Fig 2A**), confirming minimal identity shared between the control and experimental samples and providing confidence that the viruses present are not the result of environmental or reagent contamination. As an additional control, we sequenced an even mock community sample to ensure that our library preparation and sequencing techniques accurately depict microbial community composition. We found minimal differences between the expected community composition and the community composition obtained by our sequencing techniques (**Supplemental Fig 2B**).

Using methods previously outlined for quantifying virome contamination [19], we verified reduction in cellular contamination within viromes by showing a significant reduction in normalized bacterial 16S rRNA gene levels in the purified viromes compared to the unpurified whole metagenomes (**Supplemental Fig 2C**). We also supported virome purity using a previously described method [15] by mapping significantly more sequences from the virome to the whole metagenome, compared to the reverse **(Supplemental Fig 2D**). Finally, we confirmed a significant reduction of contamination from human cells in the virome, compared to the whole metagenome **(Supplemental Fig 2E**). These analyses suggest that viral reads are in greater abundance after VLP purification, and reinforce the utility of VLP purification techniques.

Skin virome composition

To examine the community membership of the skin virome, we used the viral UniProt TrEML reference database to annotate predicted open reading frames (ORFs) in the assembled viral contigs. Annotated ORFs were then subjected to a voting system that assigned taxonomy based on the most abundant ORF annotation within the contig, as described previously [20]. Some contigs had ties in taxonomic votes, which were labeled as having "multiple hits" because they are not unknown, but we cannot assign a resolved viral taxonomy with confidence. The abundance of each taxonomically identified contig was quantified as the number of unassembled reads that aligned to the contig. Read counts were normalized in order to account for differences in contig length, sequencing efficiency, and associated run variation of that overall sample, using methods previously described [20].

The majority of the identified dsDNA viral contigs belonged to the Caudovirales order (tailed bacteriophages), suggesting a higher proportion of bacteriophages among skin dsDNA virus communities than previously suggested [11] (**Supplemental Fig 3A)**. At the species level, we

observed bacteriophages of known skin inhabitants such as *Propionibacterium* phages and *Staphylococcus* phages (**Fig 2A**), and their relative abundances were significantly variable across different skin microenvironments (**Supplemental Fig 3B,C;** $p < 0.05$ Kruskal and Multiple Comparison Post Hoc Test) and occlusion status (**Supplemental Fig 3E-F**; $p < 0.05$ Kruskal-Wallis and Multiple Comparison Post Hoc Test). A large fraction of each virome contained contigs that maintained equal similarity to multiple phages, meaning they were not assignable to a single species, and were therefore annotated as "multiple hits" (**Fig 2A**). This is likely a reflection of the modular nature of bacteriophage genomes, and highlights the need for more robust reference databases for a better understanding of phage genome architecture. There was also an abundant representation of environmental phages, including *Pseudomonas* and *Bacillus* phages.

The most abundant eukaryotic virus was Human Papillomavirus (HPV), prominent in some individuals, and generally present in significantly greater relative abundance in sebaceous sites and exposed sites (**Supplemental Fig 3D,G;** $p < 0.05$ Kruskal and Multiple Comparison Post Hoc Test). Human Polyomaviruses (HPyVs) were detected in very low abundance, where only 6 samples contained any sequence mapping to known HPyV genomes, and no sample had >100 putative HPyV sequences.

3.4.2 Skin total microbial community composition

In addition to examining the taxonomic composition of the virome, we further characterized the membership of the whole microbial skin community using the corresponding sample set that was not subjected to VLP or microbial selection. Bacterial communities were classified from the unassembled sequences using MetaPhlAn [21], which annotates sequences based on clade-specific markers from reference genomes. Additionally, bacterial, fungal, and viral species abundances were quantified from assembled contigs using the lowest common ancestor algorithm

in MEGAN [22]. Consistent with previous whole metagenome analyses of skin [2, 23], *Propionibacterium* (including *P. acnes)*, *Staphylococcus* (including *S. epidermidis* and *S. hominis)*, and *Corynebacterium* were the dominant bacterial genera (**Fig 2B and Supplemental Fig 4A,B**) and *Malassezia* was the most abundant fungal genera (**Supplemental Fig 4A,C**). Viruses were present in low abundance (average 0.4% per sample), likely due to the relatively small genome size of viruses compared to prokaryotes and micro-eukaryotes, and further highlights the utility of VLP isolation before sequencing (**Supplemental Fig 4A,D**). The viruses recovered were primarily "unclassified" and *Staphylococcus* phages (**Supplemental Fig 4D**).

3.4.3 Variation of the skin virome and total metagenome among anatomic sites

As demonstrated above, and extensively in previous literature [15-18], the majority of viruses were taxonomically unidentifiable due to insufficient reference database information. In order to capture information from both characterized and uncharacterized genomes, we employed reference-independent approaches based on relative abundance of each contig in our dataset. To assess the beta diversity ("between sample diversity") among anatomical sites, we calculated the Bray-Curtis dissimilarities between communities at the same and different anatomical sites. We identified significant differences in virome and whole metagenome community structure based on microenvironment and occlusion status (**Fig 2C,D**; Adonis test; $p < 0.001$). These findings parallel previous reports of the bacterial and fungal skin microbiomes [24, 25], and highlight an additional role for occlusion/exposure parameters in microbial community structure and function.

We further estimated and compared alpha ("within sample") diversity of viral communities using a reference-independent approach of calculating the Shannon diversity index. Here we estimated virome diversity, including the viral dark matter, using the PHACCS toolkit [26], which calculates the degree of contig assembly to generate a "contig spectrum" that is compared to

simulated communities of varying size and diversity until a suitable match is found. PHACCS predicts the virome size and diversity as if the entire community (both known and unknown viruses) were sequenced and annotated. Shannon diversity of bacterial communities among anatomical sites was calculated based on reference-dependent taxonomic relative abundance information described above. We found that the virome and bacterial metagenome of sebaceous sites was less diverse than moist or intermittently moist sites (**Fig 2E-G**; $p < 0.05$ Kruskal and Multiple Comparison Post Hoc Test). While the virome was most diverse at intermittently occluded sites (e.g. Ac), the bacterial metagenome was most diverse at occluded sites (e.g. Tw and Um; **Fig 2E-G;** $p < 0.05$ Kruskal and Multiple Comparison Post Hoc Test), further highlighting the differences in viral and bacterial community diversity based on anatomic sites.

To assess the utility of reference-independent methods in determining differences in viral diversity, including that of the viral dark matter, we performed the above alpha and beta diversity analyses using the reference-dependent taxonomic relative abundance information from Figure 2A. Alpha diversity of the reference-dependent dataset (**Supplemental Fig 5A-B**) was strikingly less than that predicted by the reference independent methods employed by PHACCS (**Figure 2F**). In contrast to the PHACCS-based analysis, there was no significant difference between the microenvironment or occlusion categories using the reference-dependent data. Beta diversity between sites of different microenvironment and occlusion status mirrored reference-independent findings (**Figure 2C and Supplemental Fig 5C-D**). Therefore, there is added value to using viral dark matter in some community analyses, but some metrics can be performed effectively with reference-based approaches.

3.4.4 Variation of the skin virome and whole metagenome over time

Previous studies suggest that temporal variation of the bacterial microbiome at a given skin site is minimal when compared to interpersonal variability [25, 27, 28], so we examined both viral and whole microbial community changes over a one-month period. There was a significant difference in shared diversity between the two time points of the viromes, but not the whole metagenomes, as measured by Bray-Curtis dissimilarity (**Supplemental Fig 6 A-B;** Adonis test; $p < 0.001$ and $p = 0.978$, respectively). These findings suggest that the whole metagenome is more stable over time than viral communities.

Using the same metric, virome temporal variability at a given skin site was significantly lower than interpersonal variability (**Fig 2H**; t-test $p = 1.26 \times 10^{-11}$), similar to what we observed for the whole metagenome (**Fig 2I**; t-test; $p = 3.50 \times 10^{-30}$). Analogous to human fecal viromes, the largest source of skin virome variance appears to be interpersonal variation [15, 16]. In contrast to the gut, which has been suggested to share over 80% of the intrapersonal virome over time [15, 16], we found less than 50% of the intrapersonal skin virome was shared over time (**Supplemental Fig 6C**).

3.4.5 Evidence of a temperate replication style

Bacteriophages can exist as lytic or temperate phages. Lytic phages lyse the host soon after infection and do not exist in a latent, lysogenic state. Conversely, temperate phages are able to integrate their genomes into the bacterial host genome and exist as prophages, as well as excise and go through the lytic cycle. To examine the replication strategies of the phages residing on the skin, we used an established approach [16] of searching VLP contigs for temperate phage replication markers, including 1) the presence of integrase genes, 2) the presence of temperate prophage genes, and 3) nucleotide identity to bacterial genomes indicative of integration. Of the 6,661 contigs that were annotated as bacteriophages by our taxonomic criteria above, 5,363 had at

least one of these three temperate phage markers (**Fig 3A**). More specifically, 592 (8.8%) contained at least a single integrase gene as represented in the UniProt TrEMBL database, 856 (12.9%) aligned to known bacterial genomes, including Actinobacteria, Firmicutes, and Proteobacteria, and 5,137 (77.1%) contained open reading frames (ORFs) similar to annotated prophage genes found in the ACLAME database of mobile genetic elements [29]. By these measures, each anatomical skin site had a median relative proportion of > 85% temperate phages, with different relative abundances by site (**Fig 3B**; $p < 0.05$ Kruskal and Multiple Comparison Post Hoc Test). This data suggests that the majority of identifiable Caudovirales bacteriophages on the skin are temperate, consistent with studies of the human gut virome [15, 16].

3.4.6 Virome functional potential and auxiliary metabolic genes

Though our data support a lesser role for host lysis in skin dsDNA bacteriophage populations, they likely influence bacterial communities via prophage integration and genetic exchange. We therefore investigated the genetic functional potential of skin viral communities compared to the whole metagenome. Functional pathways were interrogated by comparison to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [30] and analyzed using the HUMAnN annotation and quantification program [31]. Overall, the virome was enriched in information processing and peptide transport, while the whole metagenome was enriched for metabolic processes (**Supplemental Fig 7A**). Gene ontology (GO) analysis revealed significant enrichment of viral components and processes, DNA transcription, and RNA metabolic processes in the virome (**Supplemental Fig 7B**), while the whole metagenome was enriched in cellular nitrogen compound and carbohydrate derivative metabolic processes. Notably, the virome was significantly enriched in the GO term "establishment of viral latency" **(Supplemental Fig 7B),** consistent with the observed dominance of temperate phages on the skin.

Some bacteriophages are known to encode auxiliary metabolic genes (AMGs; "host" genes within phage genomes) that promote viral infection by modulating host metabolic activity (reviewed in [13]). We evaluated whether there were core AMGs conserved across the entire skin virome, thereby belonging to the overall core gene set. To accomplish this, we clustered the predicted virome contig ORFs into representative OTU-like sequences called Operational Protein Families (OPFs) [32, 33]. Core OPFs were defined as those OPFs that were present in all samples from a skin site. Core OPFs were differentially distributed across skin sites, with the greatest amount present on the forehead (**Fig 3C**). Of the 15 core OPFs present in all virome samples, all were hypothetical or known phage genes, and none were AMG candidates (**Fig 3D**), suggesting a sparse population of core skin virome AMGs. As highlighted above, in comparison to the metagenome, the virome was enriched for KEGG pathways related to transport (**Supplemental Data Fig 7B**), as well as GO-terms associated with regulation of RNA metabolic processes (GOEast, p-value<0.05). While not strictly belonging to a "core" set of genes, this indicates that potential AMGs are present throughout the skin virome. We also investigated the distribution of OPFs with respect to skin site microenvironment and occlusion and found significant differences (Bray-Curtis dissimilarity; adonis test; $p<0.001$), suggesting differential spatial distribution of virome functional potential (**Fig 3E**).

3.4.6 Antibiotic resistance and virulence factor enrichment

Because phages may alter the phenotypes of their hosts by conferring novel virulence and pathogenicity functions, we investigated the potential for antibiotic resistance and bacterial virulence encoded within the skin virome. Using blast algorithm parameters specified in previous foundational human virome studies [16, 34], we assessed antibiotic resistance potential by comparing ORFs from the assembled virome contigs to the Comprehensive Antibiotic Resistance

Database (CARD) [35] (blastx; e < 10e-^{5}). To further increase our confidence in the annotations beyond that of past studies, we filtered the blastx hits to keep only those with > 75% identity. Viromes contained 29 unique antibiotic resistance gene (ARG) groups, which were related to antibiotic efflux, and resistance to beta-lactamases, rifampin, tetracycline, and elfamycin (**Fig 4A**). Tetracyclines are commonly used to treat dermatological conditions such as acne, and elfamycins are naturally occurring antibiotics with strong activity against *Propionibacterium acnes* [36]. To confirm the identified ARGs are associated with the virome and not cellular contamination or artifacts, we demonstrated ~50% of ARGs co-localized on contigs with other annotated phage genes, or are themselves known phage-associated antibiotic resistance genes (**Fig 4B**). ARGs were primarily associated with "multiple hit", *Bacillus*, and *Streptococcus* phages (**Fig 4B**). We also identified potential virulence factors (VFs) associated with the skin virome using the Virulence Factor Database (VFDB) [37] with the same blastx parameters and filtering as described for antibiotic resistance analysis above. We identified 122 unique VF genes and >1/3 of the VF contigs were either known phage-associated genes or co-localized with phage genes (**Fig 4C**). These findings together indicate that bacteriophages of the skin microbiome may be a significant source of transmissible genes associated with antibiotic resistance, virulence, and pathogenicity.

3.4.7 Inference of phage-bacteria interactions: Co-occurrence network analysis

To predict phage-bacteria interactions of the skin, we constructed a correlation network from relative abundances of bacteria and known phages, as previously described [38] (**Fig 5A**). Positive interactions indicate that the bacteria and phage typically co-occur, while negative interactions suggest a mutually exclusive relationship between the bacteria and phage relative abundance. The resulting network of significant phage-bacteria interactions contained 21 nodes, 7 bacterial and 14 phage. *Propionibacterium* and *Staphylococcus* bacteria were typically co-present

with their phage counterparts, *Propionibacterium* phage and *Staphylococcus* phage, respectively (**Fig 5A).** The overall co-occurrence structure suggests that the network is non-random, exhibiting scale-free properties such as short average path lengths (characteristic path length=2.781) and a node degree distribution that approximately fits a power law (R^2=0.781) [39]. Short average path lengths suggest the skin phage-bacteria community network is able to respond rapidly to perturbations [40]. The heterogeneity value (likelihood of uneven distribution of edges) of the network was 0.819, suggesting that there are fewer hubs, and indicating presence of potential "keystone" taxa in the network [41].

Hubs may be distinguished by identifying nodes of high degree. In the skin bacteria-phage network, *Corynebacterium*, with a degree of 10, had the greatest number of interactions, while all other nodes had a degree ≤ 5. *Corynebacterium* positively associated with eight phage, including *Corynebacterium* phage and *Staphylococcus* phage, and negatively associated with two phage, including *Propionibacterium* phage (**Fig 5A**). These features of the network topology suggest that the skin bacteria-phage network is able to rapidly respond to perturbations, and *Corynebacterium* may act as a key hub.

3.4.8 Inference of phage-bacteria interactions: CRISPRs

CRISPRs are a form of bacterial adaptive immunity against phage predators. Spacer sequences, generally 26-72 nt in length, are captured from invading phages and integrated into the bacterial chromosome. These spacer sequences provide a genomic record of phage predators encountered by the bacteria. We detected a total of 477 unique spacer sequences, identified by 68 unique CRISPR repeats in the whole metagenomic dataset. Only 18 spacers aligned to VLP contigs (**Fig 5B**). These spacers were found in 21 metagenomic contigs and mapped to 40 unique VLP contigs. Spacers found in the Um only aligned to Um VLP contigs. Two *Staphylococcus* spacers

detected in the Ax aligned to 16 different VLP contigs that were found at every body site except the Pa (**Fig 5B)**. A *Propionibacterium* spacer found both in the Pa and Tw aligned to eight different VLP contigs from the Ax, Oc, Fh, and Ra (**Fig 5B**). These findings indicate that phage-host dynamics may not be restricted by anatomical skin site, and spacers identified at one skin site may be restricting phage during invasions from other skin sites, which could in part explain spatial partitioning of the skin virome. We further characterized the genomic CRISPR targets within the VLP contigs and found that the majority of targets within coding regions belonged to phage portal proteins, which are genes involved in packaging DNA into phage particles (**Fig 5C**). Interestingly, the majority of CRISPR targets did not map to predicted ORFs, suggesting that there is not a targeting preference for genomic coding regions (**Fig 5C**).

3.5 Discussion

In summary, we present parallel analyses of the human skin virome (as determined from purified VLPs) and whole metagenome. Purification of VLPs provides many advantages for virome-targeted analyses, including deeper sequencing of viruses and the ability to confidently assess viral dark matter using reference-dependent and –independent approaches. However, this technique has previously been technically prohibitive for application to skin viruses, due to low amounts of microbial burden in and on the skin. Advanced library preparation techniques utilizing ultra-low amounts (<1 ng) of DNA have facilitated this study to characterize the human skin dsDNA virome in parallel with the whole metagenome in order to gain insight into multi-kingdom interactions of the skin microbiome.

Our results demonstrate that the skin virome is highly site specific, and is modulated by occlusion and exposure, in addition to sebum and moisture. This significant effect of skin occlusion on viral and whole microbial communities has not yet been described in previous skin whole microbial

analyses, and provides new insight into the variation of these communities across anatomical sites. Anatomical intrapersonal and interpersonal variation play a greater role in cutaneous viral community composition than intrapersonal temporal variation, supporting the role for persistent commensal populations, rather than a dominance of new acquisition of different transient viruses from the environment. Though our study provides some insight into the temporal dynamics of skin viral communities, a limitation is that our time series consisted of just two time points separated by one month. Similar to studies of the bacterial microbiome, high frequency temporal virome analyses are needed to further improve our understanding of skin virome dynamics and the degree to which temporal variability is a personalized feature [42].

The persistence of phage populations on the skin, and especially dsDNA phages, is possibly due in part to the temperate nature of their infections. While cutaneous phages that are primarily temperate may not exhibit a predator-prey dynamic with their hosts, they may give rise to novel bacterial strains via transfer of genes including antibiotic resistance and virulence factor genes, which were found in our samples. The dynamics of phage predator/prey relationships within communities is complex, and while our study provides a first look into these community dynamics in the skin, further studies will be needed to more completely characterize these relationships.

Although we noted that the majority of identifiable phages in the sampled skin virome were temperate, we were only able to predict the replication styles of the identifiable phages. This highlights the need for robust reference databases and the utility of reference independent methods. Additionally, we were not able to detect ssDNA viruses or enveloped viruses. Because of our efforts to confirm a reduction of bacterial genomic DNA in our samples, we are confident

that the majority of the sequences are in fact from free phages, and provide a valuable description of our identifiable virome library.

Bacterial hosts corresponding to some of the most abundant phages, including *Pseudomonas* and *Bacillus* phages, were not abundant in the skin bacterial communities. Because *Pseudomonas* and *Bacillus* bacteria are common environmental inhabitants, it is possible that their phage predators are in fact persistent, transient colonizers that immigrate onto the skin from their ubiquitous external sources, rather than commensal inhabitants of the skin.

In addition to showing complex community dynamics within the skin viral communities, we also provided evidence for potential interactions between the virome and the other microbial communities using co-occurrence network modeling and CRISPR identification techniques. Our network analysis allowed us to infer an extensive and multikingdom ecosystem structure. Understanding these ecological interactions, and experimentally validating them, will be critical for further developing targeted therapeutics such as phage therapy. One potential limitation of these analyses is that inferences relied upon sampling from contralateral sites. While contralateral skin sites may not be identical in microbiome composition, they have previously been shown to be very similar [27, 28, 43, 44].

CRISPR analysis suggested differing degrees of ongoing phage infections at different sites, or simply differential abundances of CRISPR arrays in the resident bacteria. CRISPRs not only targeted phages found at the same skin sites, but also targeted phages at other skin sites, providing a record of successfully repelled attacks from phages now detected at other body sites. These findings suggest a potential mechanism for partitioning of the skin virome between different anatomical locations and warrants further investigation. Finally, we found that CRISPRs most

often targeted phage portal proteins when targeting a phage coding region, but there did not appear to be a selection for CRISPRs targeting coding phage genomic regions. While we focused on CRISPR mechanisms of interaction, there are other mechanisms of bacteria-phage interactions that are worth investigating in future studies such as restriction modifications.

The viral relative abundance profiles presented in this study differ from those reported in previous whole metagenomic-based studies, which observed skin phage populations primarily dominated by *Propionibacterium* phages, *Staphylococcus* phages, Human Papillomaviruses (HPVs), and Human Polyomaviruses (HPyVs) [2, 11, 12]. In contrast, our study found prominent levels of *Pseudomonas* phages, with relatively lower levels of *Propionibacterium* phages, *Staphylococcus* phages, and HPVs, and very low levels of HPyVs. Methodological differences most likely account for these disparities. Foulongne et al used whole metagenomic techniques without virus purification and also utilized multiple displacement amplification (MDA) techniques for amplification. In addition to quantitative biases associated with amplification, MDA is biased toward ssDNA viruses (such as HPyV) and produces artifacts such as chimeras [45-47]. Whole metagenome preparation without virus purification is well-known to over-represent different phages from those identified in virome purification-based studies [48]. One reason for these biases is that virome purification methods only detect bacteriophages and other viruses that are encapsulated as VLPs. In contrast, whole metagenomic techniques detect all free DNA present in the sample, regardless of whether it is present as a free phage particle, or an integrated prophage.

This is particularly important when considering the high relative abundance of *Propionibacterium* phages observed in previous studies. *Propionibacterium* phages are thought to persist in their bacterial hosts in a pseudolysogenic, plasmid state [49, 50], meaning that while the phage does not persist by integrating into its bacterial host genome, it persists as a plasmid within the

bacterium host. Thus, the plasmid DNA would be destroyed with the bacterial host genomic DNA during virome purification, leaving only free *Propionibacterium* phage VLPs. Without such a purification step, the episomal DNA would be sequenced and considered part of the virome, even though that DNA was latent and did not exist as a VLP in the virome community. This could lead to biases in detection of other phages as well.

There are also biases associated with our methods that are important to take into account when interpreting our results. Our virome purification methods utilize chloroform/DNase treatment, which does not allow for the detection of enveloped viruses. Methods similar to those utilized in our study are also susceptible to over representation of certain phages, and can differ to various degrees compared to virome studies utilizing filtration instead of chloroform/DNase treatment or even treatment with DTT [48, 51]. Despite these caveats, we utilized viral purification techniques that are considered to be the gold standard of the field, allowing us to make meaningful parallels to studies of other human virome systems.

Overall, the findings outlined here set the stage for future studies of a) acquisition of viral communities, b) responses to perturbations such as antibiotic therapies and hygienic routines, c) factors impacting temperate vs. lytic replication cycles (i.e. DNA damaging UV radiation or antibiotics), and d) impacts on human health and disease. Long term, this work may also inform potential therapeutic strategies for skin disorders based on phage therapy.

3.6 Materials & Methods

This study represents a major body of methodological work that allowed for robust virome sample preparation and analysis, and will provide a valuable guide for future studies. We outline optimized methods for VLP DNA purification from low biomass samples, and provide rigorous

analysis outlines and source code that can be used in future analyses, even beyond the skin. Please see the supplemental methods section for a detailed description of our methods, as well as the source code and intermediate data files related to all experiments.

3.6.1 Sample collection

In short, we recruited a cohort of 16 healthy individuals (ranging from 23 to 53 years old) in accordance with our protocol approved by the University of Pennsylvania Internal Review Board. Sample collection was performed following informed consent by the subject. Exclusion criteria included self-reported antibiotic treatment (oral or systemic) six months prior to enrollment, observable dermatologic diseases, and significant comorbidities including HIV or other immunocompromised states.

3.6.2 Sample sequencing and processing

Whole metagenome DNA was prepared from cutaneous swab samples using techniques similar to those previously described [28]. The VLP DNA extraction protocol was optimized from a previously described method [14], which allowed for efficient VLP isolation and DNA purification from the low biomass samples. The DNA was prepared for sequencing using an optimized protocol for the Illumina Nextera XT library preparation kit. Sequencing was performed on the Illumina MiSeq and HiSeq2500 rapid chemistry platforms. All community analyses were performed using custom Bash, R, and Perl scripts, building off of established concepts and utilizing existing algorithms and toolkits.

Quality control was performed to remove sequencing adapters, low quality sequences, and sequences with similarity to the human genome [52]. Mock samples were also collected to control for background sequencing signals. We performed follow up analyses on this control data

to ensure a high quality sequence set. Contigs were assembled using the high quality sequences in the Ray *de novo* assembly program [53].

3.6.3 Taxonomy and Diversity

As has been previously described, virome taxonomy was assigned by annotating open reading frames using the UniProt reference database [54], and assigning contig taxonomy based on the most frequent ORF taxonomy present similarity [20]. Alpha diversity was estimated including both the known and unknown viruses using the PHACCS algorithm [55]. Beta diversity was assessed using the Bray-Curtis dissimilarity metric within the VEGAN R package (CRAN) [56], and was based on normalized sequence counts (RPKM) for each contig by sample [56]. This beta diversity information was also used for the intra- and inter-personal diversity calculations. Whole metagenome diversity was calculated using the VEGAN R package.

3.6.4 Prediction of bacteriophage replication cycle distribution

Virome replication cycle distribution was calculated by quantifying the presence of temperate marker genes, including integrase, prophage elements within the ACLAME database [29], and bacterial reference genome elements. Sequences were mapped back to the temperate and lytic contigs to assess normalized relative abundance.

3.6.5 Functional annotation and comparison

Sequence functionality was predicted by mapping reads to a reduced KEGG reference database [30] and annotating them with the HUMAnN program [31]. Gene ontology enrichment analysis was performed in GOEAST [57], using open reading frames that were predicted using the Glimmer3 toolkit [58] and blasted against the UniProt reference database. Operational protein family and auxiliary metabolic gene analysis was performed similar to previous studies [32, 33].

The Comprehensive Antibiotic Resistance Database (CARD) [35] and Virulence Factor Database (VFDB) [59] were used with predicted ORFs to estimate the potential for antibiotic resistance and virulence, respectively.

3.6.6 Inferred interactions between phage and bacteria

Inferred interactions between phages and bacteria were calculated using CoNet [60] within Cytoscape [61], as previously described [38]. Only interactions supported by two of the five tested metrics (Pearson and Spearman correlation metrics, the Mutual Information similarity metric, and Bray Curtis and Kullback-Leibler distance metrics) were retained for analysis of potential interactions.

3.6.7 CRISPR identification and comparison to the virome

CRISPR targeting of the bacterial hosts against the viruses was performed using the PilerCR program for CRISPR identification within bacterial genomes [62]. The CRISPR spacer sequences were mapped against the phage contigs from various locations to evaluate potential targeting. Phage ORFs targeted by spacers were identified using the UniProt TrEMBL database and blastx ($e<10^{-10}$).

3.7 Data Access

Sequences are deposited in the NCBI Short Read Archive under BioProject PRJNA266117 and SRA accession number SRP049645. Analysis scripts described in the Methods and intermediate files are archived at Figshare and available at DOI: 10.6084/m9.figshare.1281248 (http://figshare.com/s/e368c0088f6111e4bb9a06ec4bbcf141).

3.8 Acknowledgements

We thank the volunteers for their participation in this study, Penn Next Generation Sequencing Core for sequencing support, the Penn Medicine Academic Computing Services for computing resources, Brian Kim (Washington University) for discussions, and members of the Grice and Bushman laboratories for their underlying contributions. This work was supported by a grant from the NIH (AR060873 to EAG). GDH and AJS are supported by the Department of Defense National Defense Science and Engineering Graduate fellowship program, JSM is supported by NIH T32 HG000046 Computational Genomics Training Grant, and BPH was supported by NIH T32 AR007465 Dermatology Research Training Grant.

3.9 Author Contributions

GDH and EAG conceived and designed the study. AST collected skin swabs from subjects. GDH, JSM, AJS, and AST prepared samples for sequencing. GDH, JSM, BPH, and QZ analyzed sequence data. SM and FDB contributed software and help with analysis. GDH, JSM and EAG drafted the manuscript.

3.10 Disclosure Declaration

The authors declare no conflicts of interest.

3.11 Figures

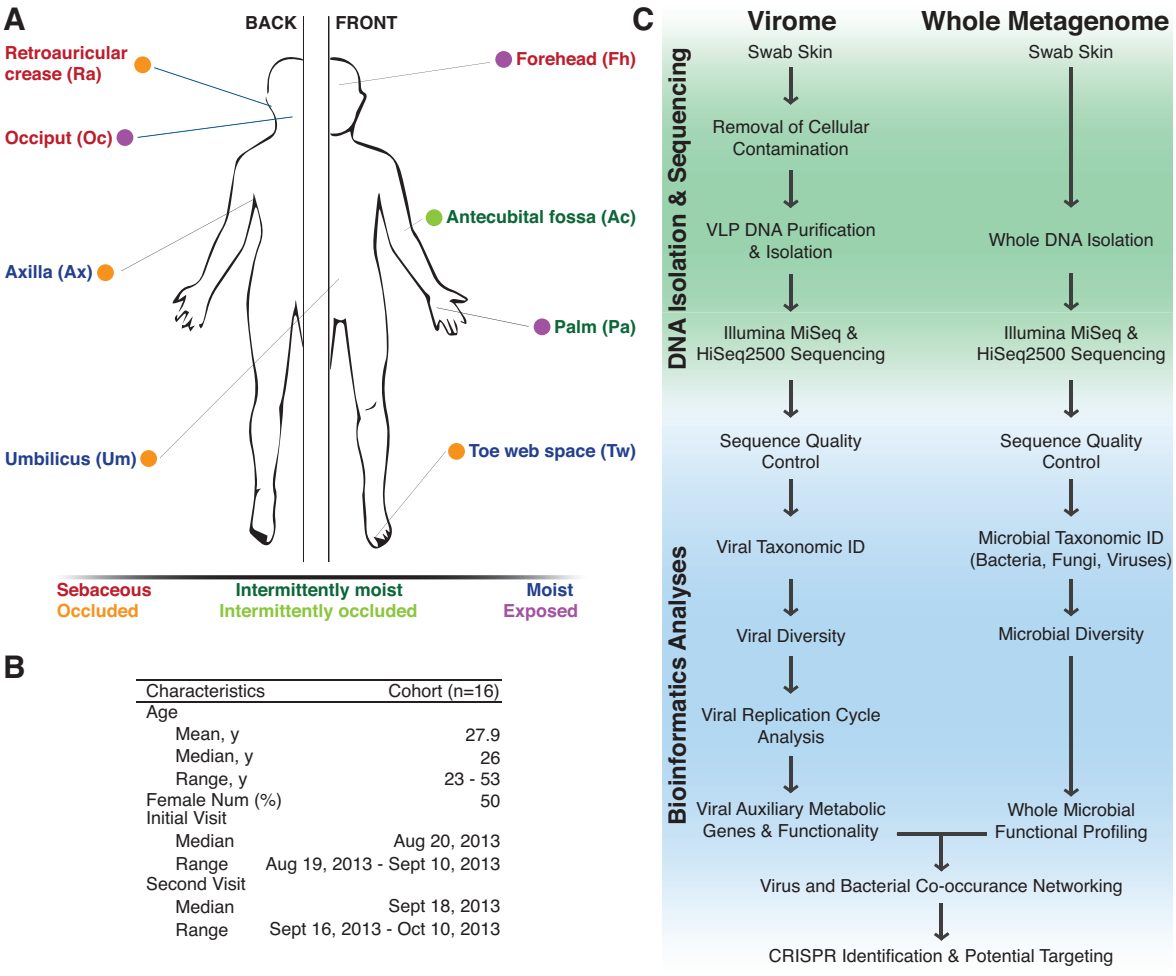

Figure 1 Study design for analyzing cutaneous viral and whole metagenomic communities.

(A) Eight skin sites were sampled from 16 subjects. Colored text indicates the microenvironment classification and colored ball represent occlusion status of the anatomical site. (B) Characteristics of the sampled cohort. (C) Flowchart illustrating procedures by which DNA was isolated from cutaneous swabs and sequenced for downstream bioinformatics analyses.

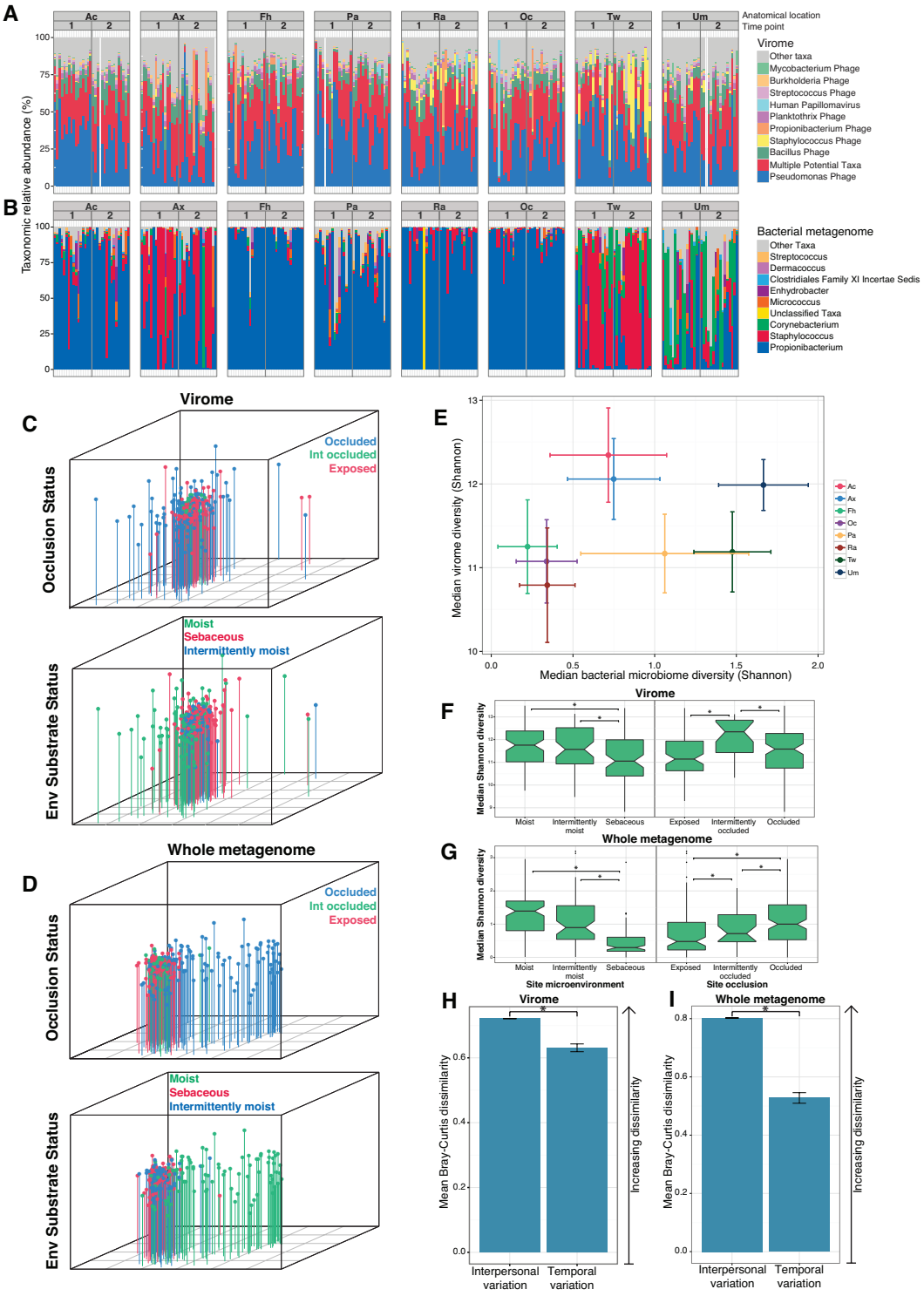

Figure 2 Taxonomy and diversity of cutaneous viral and bacterial metagenomic communities

(A,B) Taxonomic relative abundance of the viral (A) and bacterial (B) communities by site over time. The viral relative abundance plots show the 10 most abundant taxa according to virus TrEMBL annotated contigs. The bacterial communities show the 10 most abundant taxa according to MetaPhlAn analysis. Each bar represents a single sample from a subject, with the bars separated by time point and anatomical location, as labeled near the top. (C,D) NMDS ordination plots of Bray-Curtis dissimilarities between virome (C) and whole metagenome (D) samples, show significant clustering ($p < 0.001$; adonis) by occlusion status and environmental substrate. (E) Alpha diversity (Shannon diversity metric) of the virome and bacterial metagenome for each anatomical site. The x-axis represents median bacterial metagenome diversity and the y-axis represents median virome diversity. Each point is the median diversity for the two communities, and error bars indicate the population notch deviation (PND) of the median. (F,G) Viral (F) and microbial (G) Shannon diversity is presented by site microenvironment and occlusion, with asterisks (*) indicating significance of $p < 0.05$ using the Kruskal-Wallis and multiple comparison post hoc test. Boxplots were calculated using the ggplot2 R package. (H,I) Intrapersonal variance compared to temporal variance of virome (D) and whole metagenome (E) as calculated by mean (± SEM) Bray-Curtis dissimilarity metric. A higher value indicates higher dissimilarity. Asterisk (*) indicates significance with $p < 1.0 e -10$.

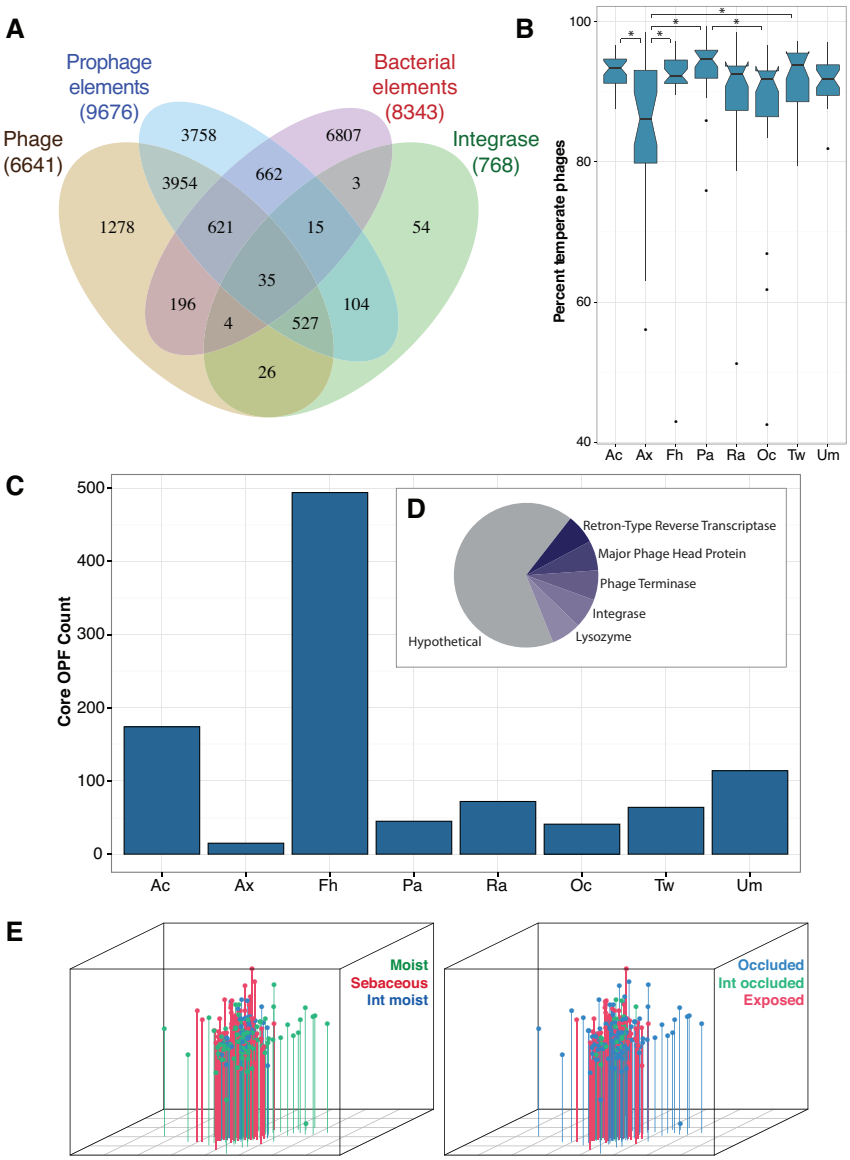

Figure 3 Replication cycle and functional enrichment of bacteriophages on the skin

(A) Euler diagram of the phage contigs (yellow) that also contain an integrase gene (green), at least one prophage element per 10kb (blue), homology to a known bacterial genome (red), or a combination of these markers. (B) Boxplot illustrating the percent relative abundance of predicted temperate phages per body site. Temperate phage contigs were defined as those contigs that contained both a phage gene at least every 10kb, as well as one of the other three temperate markers. Relative abundance was calculated as the relative RPKM of unassembled reads that

mapped back to the assembled contigs. Asterisk (*) indicates significance of $p < 0.05$ using the Kruskal-Wallis and multiple comparison post hoc test. (C) The distribution of exclusive operational protein families (OPFs) associated with each anatomical site. (D) The distribution and UniProt annotation of the 15 core OPFs found across the entire virome. (E) Bray-Curtis dissimilarity of the virome samples by OPF relative abundance. Clustering was significant ($p < 0.001$) by the adonis test for both environmental substrate and occlusion.

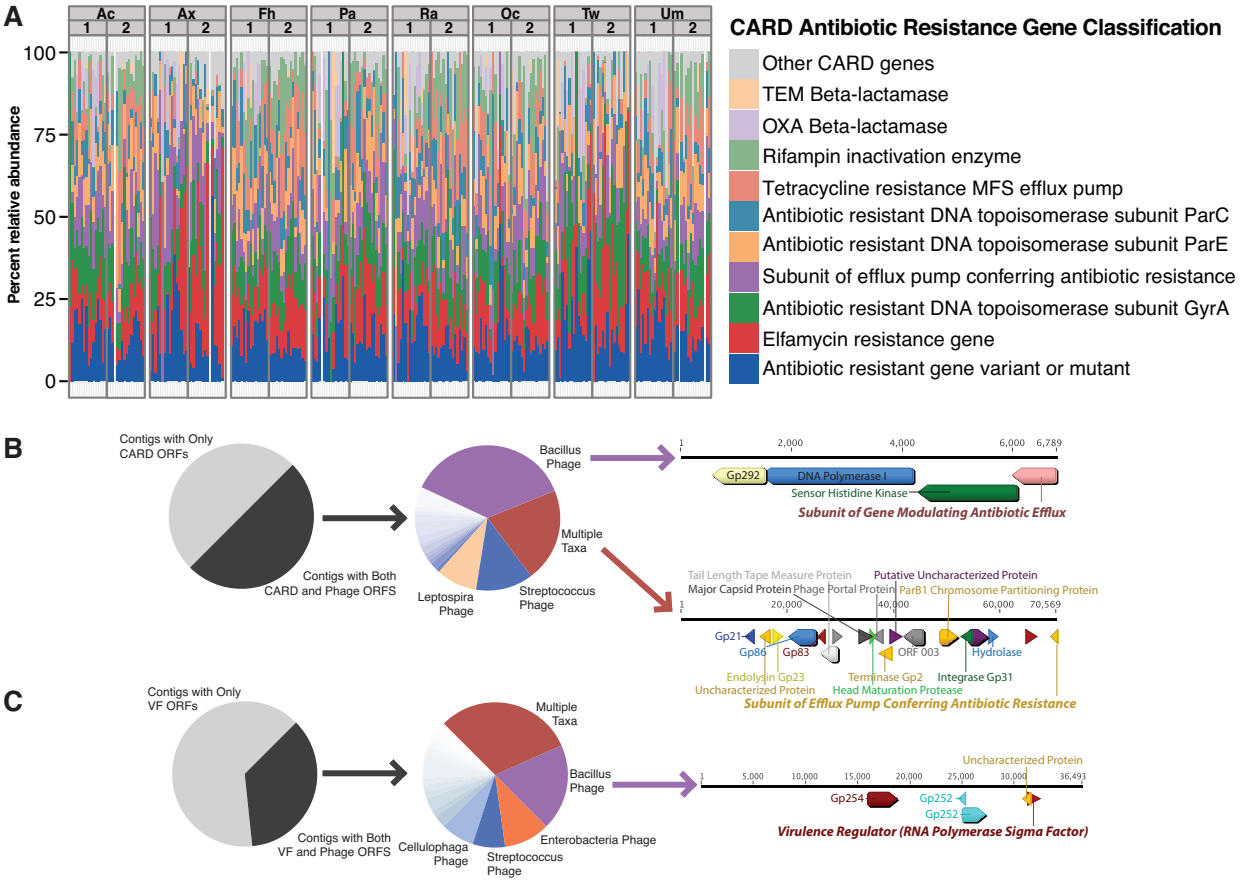

Figure 4 Antibiotic resistance and bacterial virulence in the skin virome

(A) Relative abundance of predicted antibiotic resistance genes (ARGs), according to the Comprehensive Antibiotic Resistance Database (CARD). Each bar represents a subject, with the bars separated by time point and anatomical location, as labeled near the top. (B) Flow diagram of the ARGs associated with bacteriophage contigs. First panel is proportion of ARGs that co-localize on contigs with other phage genes, or are themselves known phage-associated genes. The middle panel shows the distribution of phage taxa that contain predicted ARGs. The rightmost panel shows two annotated examples of ARGs co-localized on phage contigs, with the CARD-predicted ARGs in bold italics. (C) Similar to B, a flow diagram of the virulence factors (VFs) associated with phages. As in B, the leftmost panel shows the distribution of predicted VFs associated with phages, middle panel shows the taxonomic distribution of those phages, and the rightmost panel shows an annotated example.

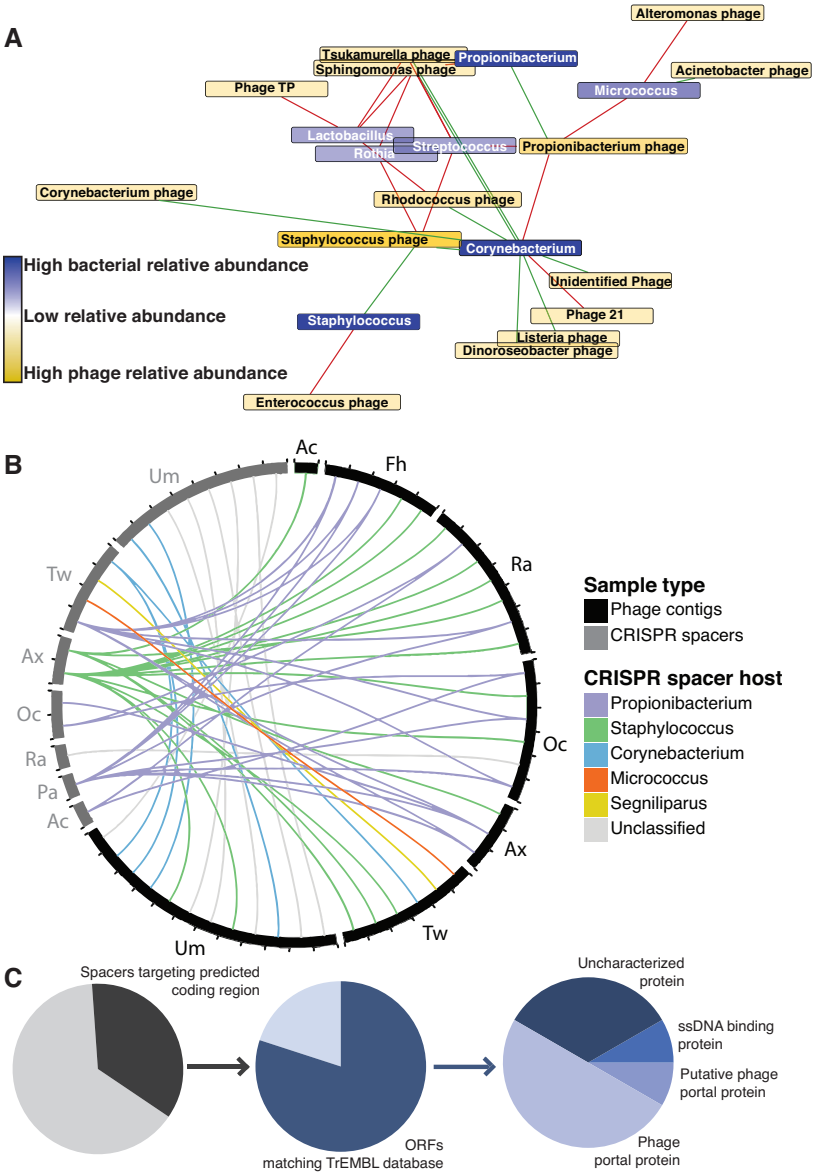

Figure 5 Modeled bacteriophage-host co-occurrence associations and CRISPR targets within the skin virome

(A) Network analysis of the correlations between bacteriophages of the virome and bacteria of the whole metagenome. Bacteriophages are represented as yellow boxes, while the bacterial genera are represented by blue boxes. The intensity of the box color indicates the overall relative abundance of the taxa. The red edges represent a negative correlation and the green edges represent a positive correlation. (B) Radial table showing bacterial CRISPR spacers (grey) that

target viral phage contigs (black). The line colors represent the CRISPR spacer bacterial host. (C) Flow chart depicting the phage genome regions targeted by skin bacterial CRISPRs. The leftmost panel shows the abundance of spacers that target a predicted coding region (ORF) within the phage genomes. The middle panel is the distribution of ORFs matching a gene in the TrEMBL reference database. The rightmost panel is the distribution of annotated coding region CRISPR targets.

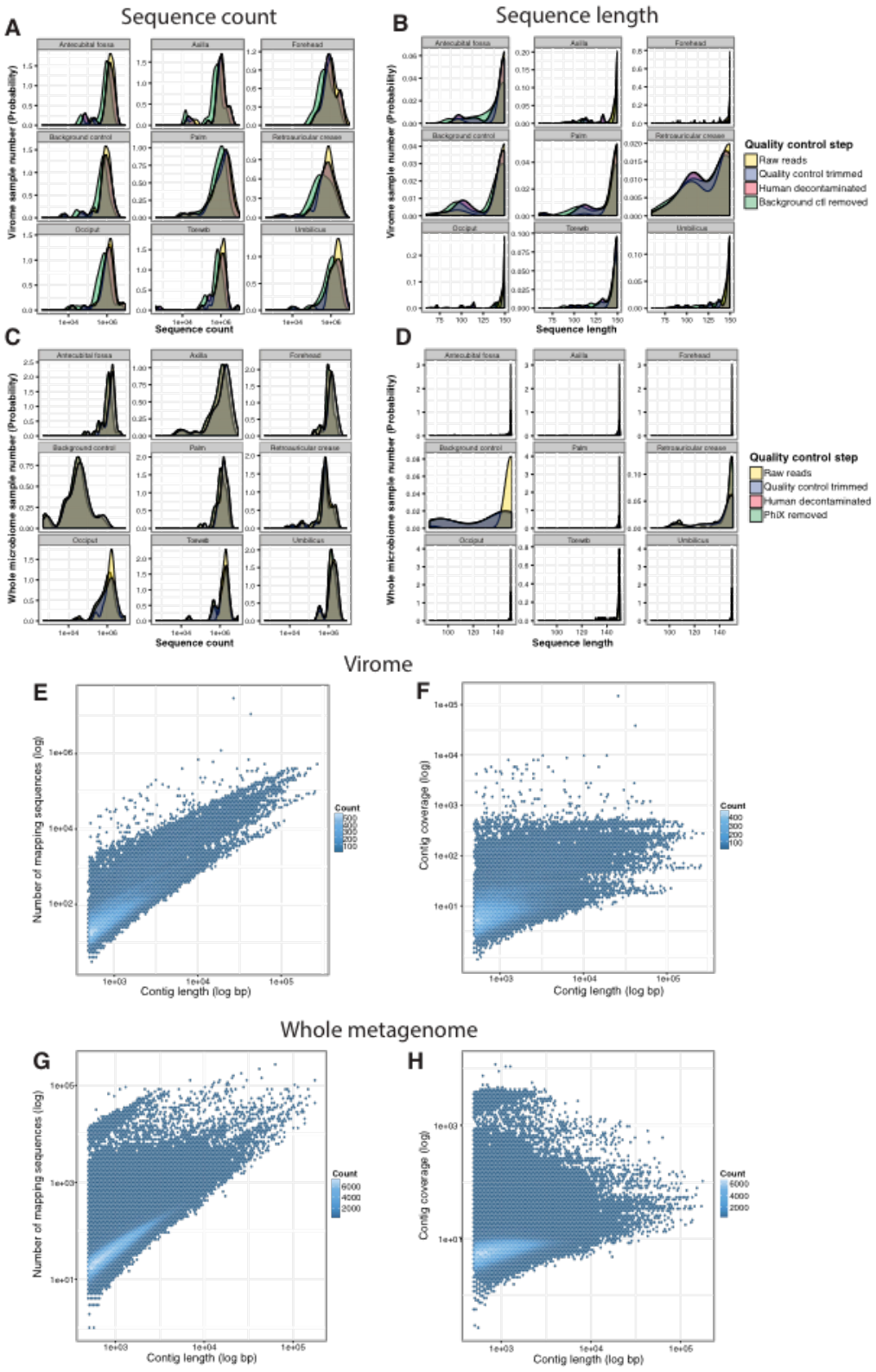

Figure S1 Contig coverage, counts, and length

A) Density histogram of each sample's sequence count and B) median sequence length, grouped by skin site. The original, un-filtered raw sequence density is colored yellow, the sequence density following quality trimming is colored blue, human decontaminated sequence density (which followed quality trimming) is colored in red, and background control cleaned densities (which followed human filtering) is green. C-D) Whole metagenome sequence statistics with the same format as A-B, except PhiX was removed instead of the background control. E) The number of unassembled reads and F) genomic coverage of the assembled virome and metagenome (G-H) contigs plotted against contig length as a contour scatter plot. Dark blue indicates lower numbers of mapped sequences or coverage, while white indicates high sequence mapping or coverage.

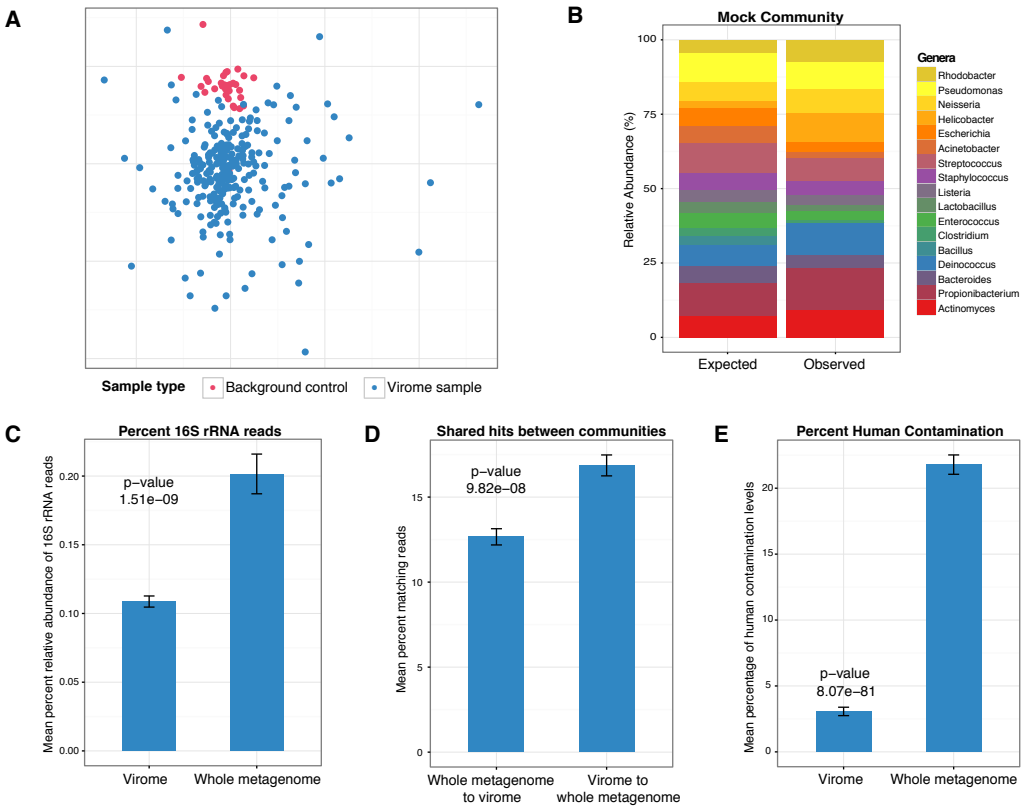

Figure S2 Quality control

A) NMDS ordination plot of Bray-Curtis dissimilarity between the background control samples (red) and experimental virome samples (blue). The clustering of the background control samples was significantly different from the experimental virome samples (Adonis test; $p < 0.001$). B) Expected relative abundances of genera in the even bacterial mock community sample, compared to observed relative abundances from library preparation and sequencing of the mock community sample. The similarity of the two profiles validates the accuracy of the sample preparation techniques used, such as the increase in PCR cycle number to overcome the low bacterial biomass of skin samples. C) Percent of reads mapping to 16S rRNA bacterial genes in the virome and whole metagenome. There were significantly fewer 16S rRNA gene reads in the virome compared to the whole metagenome. D) The percent of whole metagenome sequences mapping to the corresponding virome libraries (blastn, e-value < 1e-5), and vice versa. Significantly more sequences of the virome mapped back to the whole metagenome, than the whole metagenome

mapped to the virome. E) The average percentage of human contamination in the virome and whole metagenome datasets. There were significantly fewer reads matching the human reference genome in the virome compared to the whole metagenome.

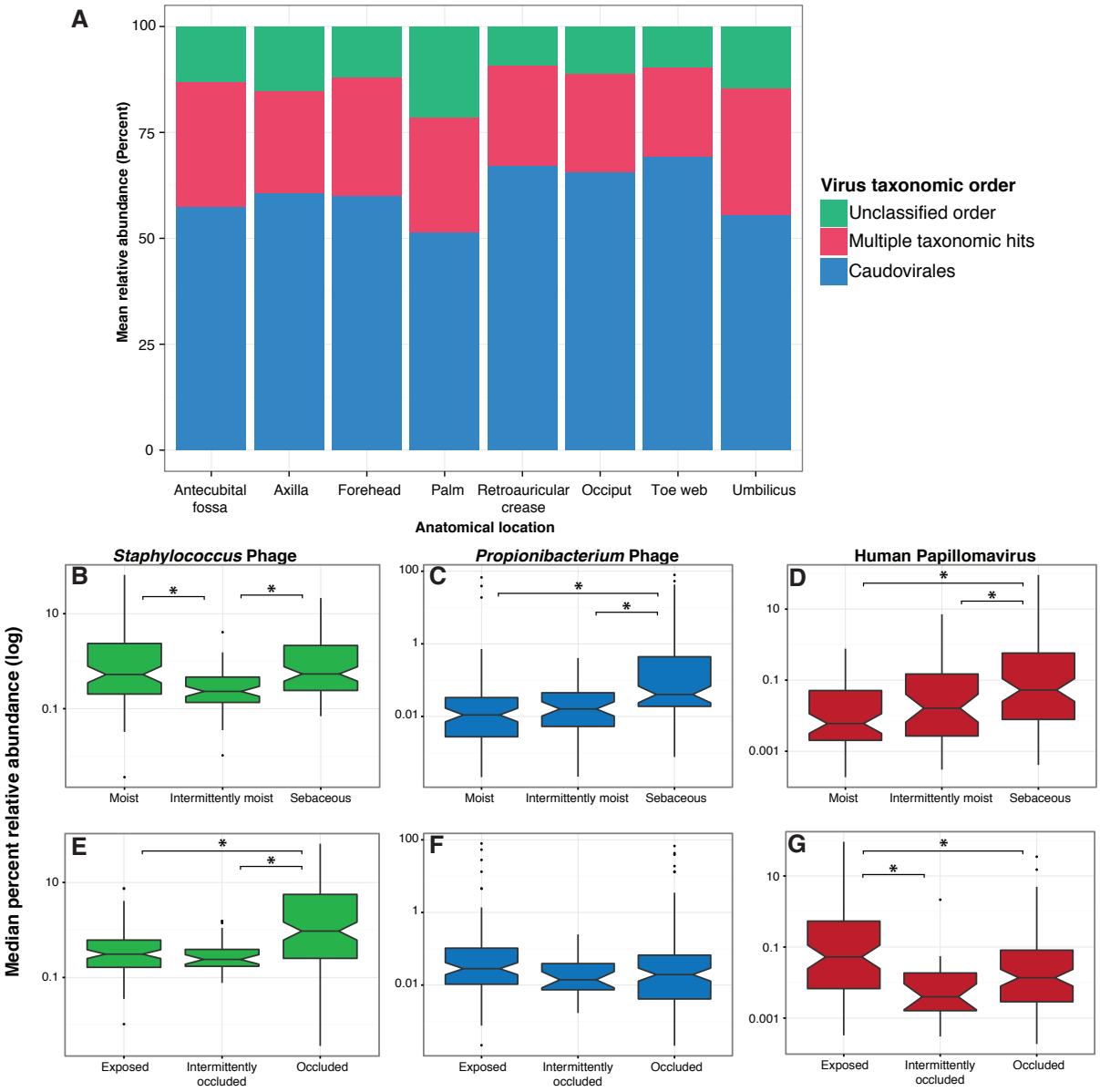

Figure S3 Classification of VLPs

A) Taxonomic order classification of VLPs, categorized by anatomic skin location. Relative abundance based on quantification of un-assembled reads mapping to annotated contigs. Only those taxa with a greater mean relative abundance of greater than 0.5% are shown. The multiple taxonomic hit classification (red) designates those reads mapping to contigs with multiple potential taxonomic identification, based on the voting-based classification scheme. Unclassified order (green) designates those reads mapping to contigs whose taxonomy has not yet been

assigned at the order level, despite specific classification at other taxonomic levels. Relative abundance of Staphylococcus phages (B,E; green), Propionibacterium phages (C, F; blue) and Human Papillomavirus (D, G; red) by site microenvironment (B-D) and occlusion status (E-G). Asterisk (*) indicates significance of $p < 0.05$ using the Kruskal-Wallis and multiple comparison post hoc test.

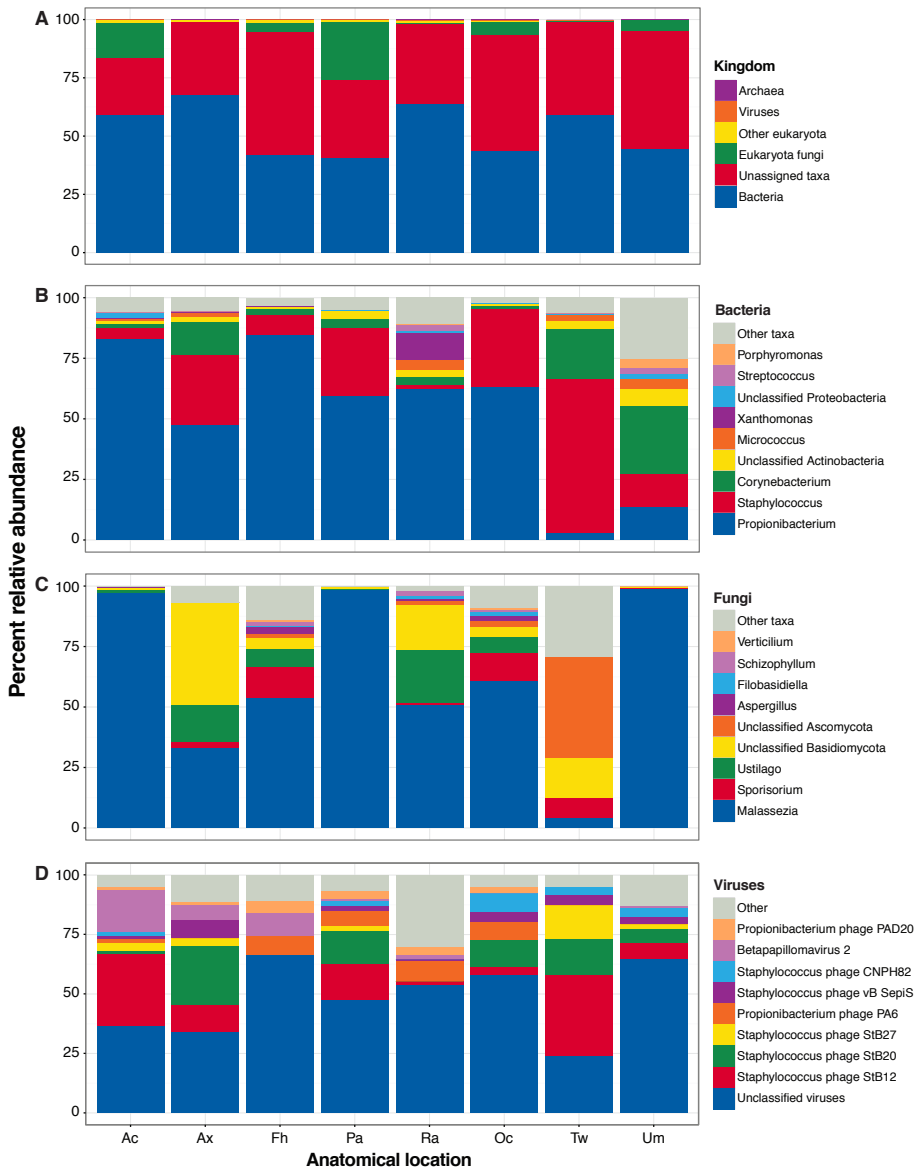

Figure S4 Multi-kingdom level classification of metagenomic sequence reads

MEGAN was used to calculate relative abundance of (A) kingdom level microorganisms, (B) bacterial genera, (C) fungal genera, and (D) viral species by skin site.

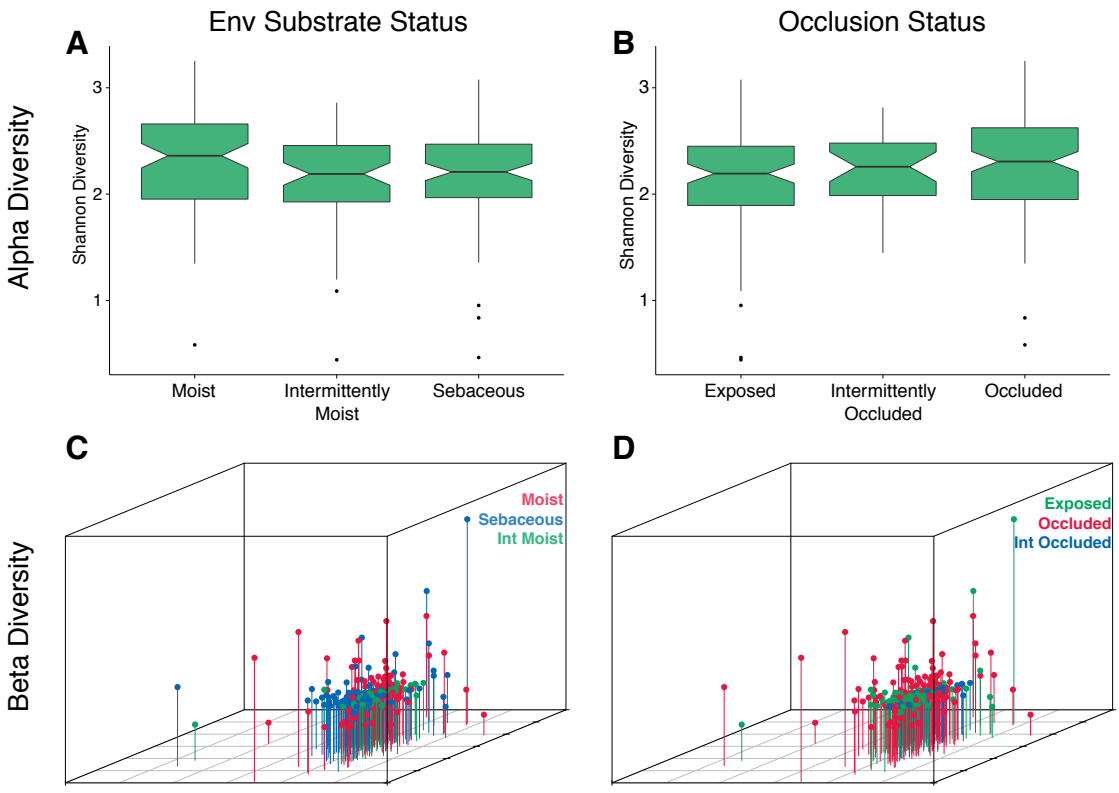

Figure S5 Reference dependent viral diversity by skin microenvironment and occlusion status

The taxonomic relative abundance information was used to calculate the diversity of the viral communities and measure their differences by skin microenvironment and occlusion status. Virome Shannon diversity was calculated using the R Vegan package, and differences were calculated using the Kruskal-Wallis and multiple comparison post hoc test. There was no significant difference between skin microenvironment (A) or skin occlusion status (B). NMDS ordination plots of Bray-Curtis dissimilarities between skin microenvironment (C) and occlusion status (D). Clustering was significant ($p < 0.001$) by the adonis test for both sample sets.

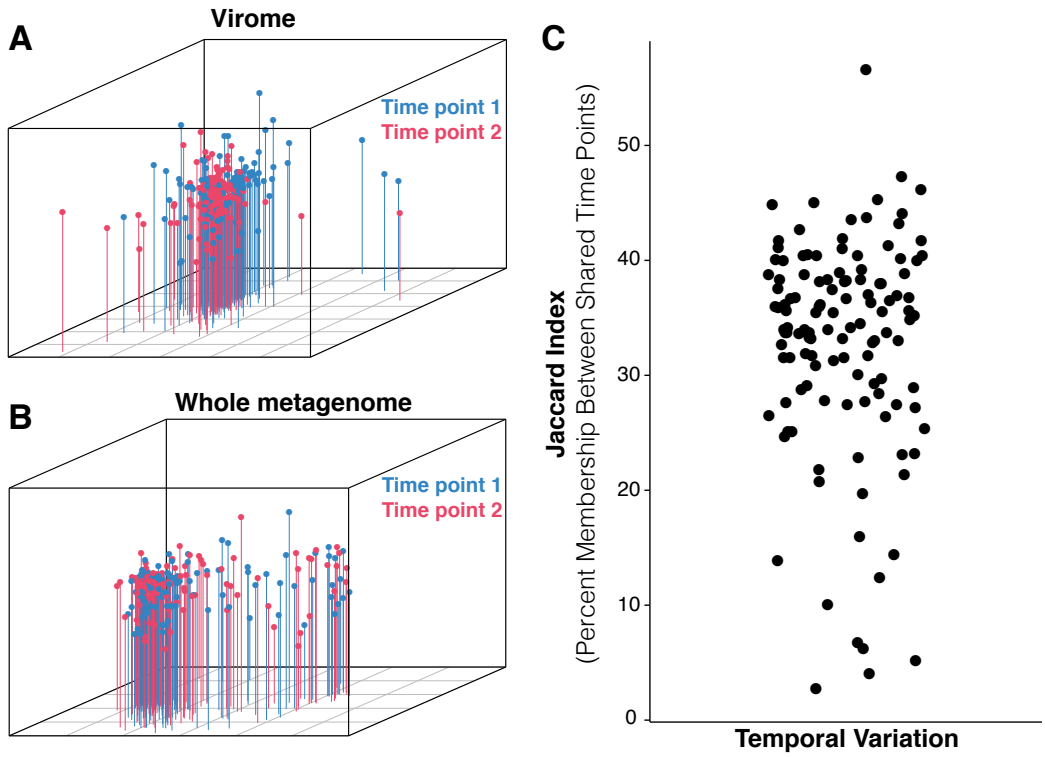

Figure S6 Temporal Bray-Curtis dissimilarities

NMDS ordination plots of Bray-Curtis dissimilarities between virome (A) and whole metagenome (B) samples, labeled by time point. Clustering was significant ($p < 0.001$) by the adonis test for the virome, but not the whole metagenome. C) Jaccard similarity index of each patient site paired over the one month sampling time. The Jaccard index was calculated using the inverse of the binary dissimilarity metric as calculated by the R base statistics package.

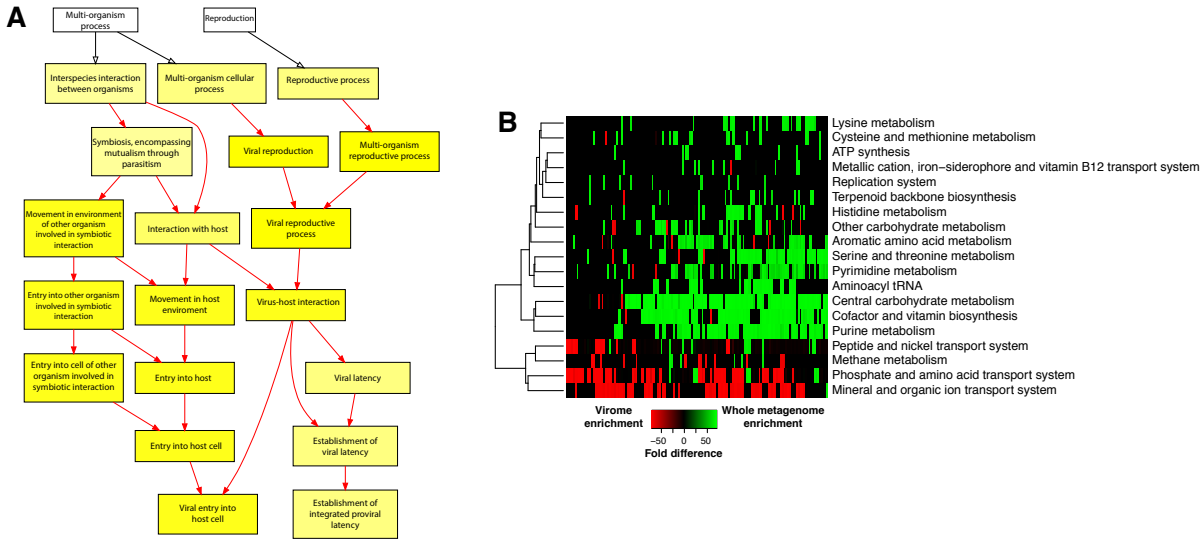

Figure S7 Functional enrichment of bacteriophages on the skin

(A) GOEAST gene enrichment analysis in sebaceous samples indicating that, compared to the whole metagenome, the virome is enriched for known viral functions including viral latency functions. Increased yellow intensity indicates a stronger significance of enrichment. Red arrows indicate relationships between enriched elements, and empty black arrows indicate a relationship between an enriched and non-enriched element. Shown is a subset of the functionally enriched GO terms under the category “Biological Processes”. (B) The heatmap depicts KEGG modules (y-axis) significantly enriched ($p < 0.05$) in skin virome (red) and skin metagenome (green). Each sample is displayed as a column across the x-axis. Dendrogram (left axis) clusters each functional group by similar enrichment profiles.

3.12 Supplemental Methods

Sample collection. The University of Pennsylvania Internal Review Board approved all human subject recruitment and sample collection. Sixteen healthy adult volunteers (ranging from 23 to 53 years old, with a median age of 26 years), residing in Philadelphia, PA and surrounding areas, both male and female (female:male ratio, 1:1), were recruited to provide cutaneous swabs from 8 anatomical locations. Subjects were swabbed at two time points, with a month in between sampling. Sample collection was performed following informed consent by the subject. Exclusion criteria included self-reported antibiotic treatment (oral or systemic) six months prior to enrollment, observable dermatologic diseases, and significant comorbidities including HIV or other immunocompromised states. Subjects were instructed to avoid hand sanitizers and antimicrobial soaps and skincare products for 1 week prior to sample collection appointment. Subjects were also instructed not to shower for 24 hours prior to sample collection appointment. Virome and whole metagenome samples were collected at the same time and at the same anatomical locations on contralateral sides of the body. Whole metagenome swabs were collected as described previously [28] and stored at -20°C immediately following collection. Virome swabs were collected using Catch-All Sample Collection Swabs (Epicentre) moistened with saline magnesium (SM) buffer (Crystalgen), and stored <12 hours at 4°C in 500µL SM Buffer. Virome swabs were prepared for short-term storage by first extracting all liquid from the swab using a DNA IQ spin basket (Promega) and centrifuging for one minute at 15,900xg. Chloroform (Fisher Scientific; 0.2 volumes) was added to the sample, gently shaken for 10 minutes, and briefly centrifuged. These samples were stored for a maximum of one month at 4°C until further processing.

Whole metagenome isolation and purification. Whole metagenome swabs were incubated for one hour at 37°C with shaking and 0.5 µL ReadyLyse Lysozyme solution (Epicentre). Samples were subjected to bead beating for ten minutes at maximum speed on a

vortex mixer with 0.5 mm glass beads (MoBio), followed by a 30 minute incubation at 65°C with shaking. Downstream isolation and purification was performed as previously described [63].

VLP isolation and purification. The VLP DNA extraction protocol was slightly modified from a previously described method [14]. Samples were centrifuged at 21,130xg for 5 minutes and the aqueous layer was transferred to a new sterile tube, with care taken to ensure no chloroform was transferred over. Samples were treated with DNase I (Invitrogen; 3 U per sample) and RNase A (Roche; 1.5 µg) for 1.5 hours at 37°C with gentle shaking to remove background host DNA. DNase I was inactivated by incubating the sample at 65°C for 10 min with gentle shaking. Virions were extracted by adding 50 µL sterile TE buffer (Fisher Scientific; pH 8.0), 5 µL 0.5 M EDTA (Gibco; pH 8.0), 500 µL formamide (Fisher Scientific), and 10 µL glycogen (0.2 mg per sample; Roche 20mg/mL) and then incubating for 30 min at room temperature. Two volumes of 100% ethanol were added and DNA was pelleted by centrifugation for 20 min at 10,000 xg and 4°C. Supernatant was removed and the pellets were washed twice with 500 µL of 70% ethanol. Pellets were re-suspended in 567 µL TE buffer (Fisher Scientific; pH 8.0) and stored at -20°C for a maximum of 1 month, until proceeding to the next step. DNA was extracted from VLPs by first removing proteins by treatment with 30 µL 10% SDS (Fisher Scientific) and 3 µL proteinase K (Roche; 20 mg/mL) followed by incubation for one hour at 55°C with gentle shaking. 100 µL 5M NaCl (Sigma) was added, mixed thoroughly, followed by addition of 80 µL CTAB (Sigma) + NaCl solution [14], gentle inversion, and incubation for 10 min at 65°C with gentle shaking. An equal volume of chloroform:isoamyl alcohol (Sigma) was added to the sample, mixed by gentle inversion, centrifuged for 5 min at 8,000 xg at room temperature, and the aqueous layer was transferred to a new tube. This process was repeated with phenol:chloroform:isoamyl alcohol (Fisher Scientific), and then again with chloroform:isoamyl alcohol. DNA was precipitated by adding 0.7 volumes of isopropanol (Fisher Scientific), gentle inversion, and incubation at either -20°C for two hours or overnight at 4°C. DNA was pelleted by

centrifugation for 15 minutes at 13,000 xg and 4°C, and the supernatant was removed. The pellet was washed with 500 µL ice cold 70% ethanol followed by centrifugation for 15 minutes at 13,000 xg and 4°C. Supernatant was gently removed, the pellets dried, resuspended in 20 µL TE buffer (Fisher Scientific; pH 8.0) and placed at -20°C for short-term storage and -80°C for long-term storage.

Virome and whole metagenome sequencing. Sequencing libraries were prepared using the NexteraXT (Illumina) library preparation kit according to the manufacturer's instructions, with the exception that PCR cycles were increased to 18 for virome samples and 15 for whole metagenome samples. Additionally, instead of using the manufacturer's NexteraXT bead-based normalization protocol, we manually normalized and pooled based on DNA concentration and average fragment lengths. Sequencing was performed at the Penn Next Generation Sequencing Core on the Illumina MiSeq and/or HiSeq2500 rapid chemistry to obtain 150 bp paired end reads.

VLP and whole metagenome sequence quality control and pre-processing (**Script P1, Script P2, Script R1**, **Script R2**)**.** Sequence data were obtained in fastq format. Adapters were removed using cutadapt (version 1.4.1) with an error rate of 0.1 and overlap of 10. Low quality sequences (quality score <33) were removed using the standalone FASTX toolkit (version 0.0.14) with default parameters. Sequences mapping to the human genome were removed from the quality-trimmed dataset using the standalone DeconSeq toolkit (version 0.4.3) with default parameters and the human reference GRCh37 [52]. Because a 1% spike-in of PhiX Control was added to the sequencing runs for quality control purposes, any sequences mapping to the PhiX174 genome (NCBI Accession: NC_001422) were also removed from the whole metagenome samples using DeconSeq.

Background Correction. We collected mock swab control samples for every subject at each time point to assess overall background contamination from either reagents and/or collection

procedures. These mock controls were prepared and sequenced exactly as the experimental samples. No significant background was recovered in whole metagenome mock controls. VLP experimental samples and mock controls were compared to the NCBI non-redundant database (downloaded October 6, 2012) using blastn in the Blast-Plus toolkit [64] (version 2.2.0) with default parameters and $e<10^{-3}$. All sample sequences whose GI-numbers matched a GI-number found in its corresponding control were removed, except for GI-numbers with only a single sequence hit. Background control samples were strongly significantly different from skin samples following subtraction (Adonis test; $p < 0.001$). In order to identify the percentage of unknown reads in the virome and metagenome datasets, reads were subsampled to 2500, blasted against the NCBI nonredundant database (blastn, $e<10^{-3}$) and input into MEGAN version 5.5.3 [65] with default parameters.

Virome Quality Assurance. To estimate reduction of bacterial contamination in the VLP dataset compared to the whole metagenome dataset, the number of sequences in each sample set matching reference reads in the GreenGenes 16S rRNA gene database (accessed July 22, 2014) were quantified using blastn ($e<10^{-5}$). Previous studies have supported the utility of VLP DNA purification methods by comparing sequence homology between the whole metagenome and VLP samples [15]. Because the viruses should be a small population in the whole metagenome samples and a dominant population in the VLP samples, a valuable VLP DNA purification protocol would result in a significantly greater number of VLP sequences matching whole metagenome sequences. Blastn ($e< 10^{-5}$) was used to quantify the number of virome sequences that match the corresponding whole metagenome sequence set, and vice versa.

Mock Community Analysis. To ensure that our library preparation and sequencing techniques were accurately depicting microbial community composition, we sequenced HM-782D Genomic DNA from Microbial Mock Community A (Even Low Concentration) in parallel with our experimental samples.

Contig Assembly. Any reads missing their corresponding paired end were removed and paired-end reads were concatenated into a single file and converted to fasta format. Contigs were assembled using the Ray Assembly toolkit [53] (version 2.3.1), using default parameters, with a minimum contig length of 500bp.

Open Reading Frame (ORF) Prediction. ORFs were predicted and extracted from contigs using the Glimmer3 toolkit (version 3.02) [58] and a minimum length threshold of 100 amino acids.

VLP sequence analysis. Taxonomic annotation and relative abundance (**Script P4, Script R4**)**.** The translated amino acid sequences of predicted ORFs from the VLP contigs were matched against a custom subset of the entire UniProt TrEMBL database that contained only virus and phage reference genes, using blastx ($e<10^{-5}$; database generation details in **Script P4**) [54]. Each contig was assigned taxonomy based on the most abundant taxa contained within that contig using a voting system as described previously for virus taxonomic assignment [20]. In brief, the voting system first annotated each ORF of a contig of interest with the best-hit virus taxonomy. It then compared all of the taxonomic assignments of the ORFs within the contig of interest, and annotated the contig with the majority ORF assignment. Contigs with less than one ORF per 10kb were not assigned taxonomy as this suggests a contig of only limited similarity [20]. Contigs without a majority ORF taxonomic assignment due to ties of multiple major taxa were assigned as having multiple possible taxonomic annotations. Because some contigs shared the same taxonomic identities, the contig table was collapsed by taxonomic identity, meaning the contig relative abundances were summed if they shared identity. Although the contigs shared taxonomic identity, we confirmed a lack of contig nucleotide redundancy by comparing all of the contigs to each other (blastn; e-value < 1e-25). No contigs mapped to any other contigs. After contigs were assigned taxonomy, taxonomic relative abundances of the contigs were calculated by mapping back all unassembled reads to the contigs using Bowtie 2 (version 2.1.0) [66] with a

seed length of 25 and one mismatch allowed per seed. The numbers of reads mapping to each contig were quantified on a per sample basis. The mapped sequence counts, contig lengths, and total sequence counts were used to normalize the sequence counts and represent the RPKM (reads per kilobase per million) of each sample to the contigs. These values were used to generate an OTU relative abundance table (each unique contig represented an OTU), which was annotated with the taxonomy described above. For more details, please see the supplemental source code and archived intermediate files. It is important to note that the definition of a bacteriophage species remains a point of active discussion, and thus a phage species was defined by the identity of a contig's predicted ORFs to an existing reference phage genome. Because reference-dependent methods of analyzing virus and phage communities can be somewhat controversial and rely on relatively small reference databases, we also use several reference-independent analyses, as described below.

Diversity analysis (**Script P5, Script R5, Script R6**)**.** Virome alpha diversity was calculated using PHACCS (PHAge Communities from Contig Spectrum; version 1.1.3) [55]. Circonspect (version 0.2.6) was used with default parameters to calculate the contig spectrum of each sample [26]. GAAS (Genome relative abundance and Average Size; version 0.17) was used to predict the average virus genome size for each sample ($e < 1e\text{-}3$, database generation details in **Script P5**)[67]. Virome beta diversity was assessed using a Bray-Curtis dissimilarity matrix calculated from the un-annotated contig OTU relative abundance table. The data were visualized in 3D using the meta-MDS (k=3) optimal clustering functionality of VEGAN (CRAN) [56], and statistical significance was assessed using the adonis test.

Intrapersonal vs Interpersonal Diversity (**Script R7**)**.** The contig relative abundance table used for Bray-Curtis dissimilarity calculation was used to calculate virome intrapersonal and interpersonal dissimilarity. Intrapersonal dissimilarity was defined as the Bray-Curtis dissimilarity measured between a specified subject's anatomical site and that same site again one

month later. Interpersonal diversity was defined as the Bray-Curtis dissimilarity measured between a specified subject's anatomical site and any other given site or subject from the same time point. Statistical significance was assessed using a two-tailed t-test. The Jaccard similarity index was used to quantify the number of intrapersonal contigs shared between time points.

Detection of human polyomavirus (**Script P6**). Unassembled VLP and whole metagenome sequences, prior to DeconSeq human sequence filtering, were queried against a custom database of reference HPyV genomes, containing 61 complete HPyV genomes from the NCBI RefSeq and GenBank databases, obtained using the search terms "*human polyomavirus* AND *complete genome*".

Whole metagenome sequence analysis. Taxonomic annotation and relative abundance (**Script P7, Script R8, Script R9**). Whole metagenome sequences were taxonomically classified using MetaPhlAn version 1.7.7 [21, 68] and MEGAN version 5.5.3 [65]. Sequences <80 nucleotides long were removed from the quality trimmed, DeconSeq filtered fastq files and one of the paired reads (SE1) was input into MetaPhlAn using default parameters. Additionally, assembled contigs were queried against the NCBI non-redundant database (blastn; $e<10^{-10}$) and output was run through MEGAN on the command-line (minSupport=5, minComplexity=0.3).

Diversity analysis (**Script R5, Script R10**)**.** Alpha diversity was calculated in VEGAN using the biom table generated from MetaPhlAn output. Beta diversity was calculated in VEGAN using the whole metagenome contig OTU relative abundance table, utilizing the same methods as applied to virome samples.

Intrapersonal vs Interpersonal Diversity (**Script R10**). This analysis was performed using the contig relative abundance table, similar to what was described above in the VLP analysis section.

Virome & bacteria metagenome diversity comparison (**Script R5**). The median virome and metagenome Shannon diversity of each anatomic site was calculated with the population notch deviations (PND). PND was calculated as PND=(1.58*IQR)/sqrt(N) where IQR is the interquartile range and N is the number of samples. This was done according to the boxplot notch calculation described in the ggplot2 R package [69], as well as in McGill *et al* [70].

Prediction of bacteriophage replication cycle distribution (**Script P8, Script R11**). Integrase protein references were collected from the UniProt TrEMBL and Swiss-Prot databases using the search terms "*organism:phage AND integrase*" (accessed data: September 02, 2014). The ACLAME database version 0.4 was used to annotate prophages [29]. Whole bacterial genomes were obtained from NCBI with the following path: <ftp://ftp.ncbi.nih.gov/genomes/Bacteria/all.fna.tar.gz>. The taxonomic summary reference information was obtained from: <ftp://ftp.ncbi.nih.gov/genomes/Bacteria/summary.txt>. Blastx ($e < 10^{-5}$) was used to identify contigs containing at least a single integrase gene, contigs containing at least a single ORF with homology to a prophage ACLAME gene for every 10kb, and contigs containing at least a single bacteriophage gene every 10kb. Blastn (>90% query length with >90% nucleotide similarity) was used to query contigs against the NCBI reference bacterial genomes. Sequences were mapped back to the temperate and lytic contigs in the relative abundance table mentioned above.

Functional annotation and comparison (**Script P9, Script R12**). One set of the paired end reads (SE1) for each sample was subsampled to 10,000 sequences, queried against a reduced KEGG reference database version 56 (blastx; max_target_seqs 1, $e < 10^{-10}$ for metagenome samples, $e < 10^{-5}$ for virome samples) [30], and input into HUMAnN[31]. ORFs subsampled at 1000 were queried against the UniProt SwissProt database (blastx; max_target_seqs 1, $e < 10^{-10}$ for metagenome samples, $e < 10^{-5}$ for virome samples) and contigs were mapped to gene ontology

IDs based on their hits. Annotated contigs were grouped by site microenvironment and input to the online Customized-GOEAST analysis tool [57] using default parameters.

Operational protein family (OPF) & auxiliary metabolic gene analysis (AMGs) (**Script P9**, **Script R13**). Functional diversity and the virome core/flexible AMGs were defined using operational protein families (OPFs; also called protein clusters). OPFs were generated by clustering predicted ORFs by sequence similarity using the UCLUST algorithm [71] in QIIME (version 1.8.0) [72] and a 75% similarity value. A representative sequence was pulled from each OPF and an OPF relative abundance table was generated by quantifying the numbers of sequences mapping to each OPF with the Bowtie2 toolkit (seed length of 25, one mismatch allowed per seed). This relative abundance table was used to predict the core and flexible OPFs, AMGs, and beta-diversity of OPFs by skin microenvironment, as described in detail in the supplemental source code. Core OPFs were defined as those that were present in all samples at a given anatomical site (i.e. core OPFs of the forehead were defined as those OPFs present in every forehead sample). Likewise, overall core ORFs were defined as those OPFs present in every sample.

Antibiotic resistance genes (ARGs) & virulence factor (VF) genes **(Script P10, Script R14).** Using blast algorithm parameters specified in previous human virome functionality studies [16, 34], we assessed antibiotic resistance potential by comparing predicted Open Reading Frames (ORFs) from the assembled virome contigs to the Comprehensive Antibiotic Resistance Database (CARD) [35] (accessed data: June 20, 2014; blastx; $e < 10e^{-5}$). To further increase our confidence in the annotations, we filtered the blastx hits to keep only those with >75% identity. Bowtie 2 (seed length of 25, one mismatch allowed per seed) was used to map all single end reads (SE1) from each sample to the CARD-annotated ORFs. The number of sequences mapping to each ORF, in addition to the total number of sequences per sample and the length of each ORF, were used to calculate RPKM values and create a relative abundance table. Numbers of contigs

containing both ARGs and bacteriophage genes were quantified, and taxonomy was assigned to the contigs containing ARGs. Contig annotation was performed using our custom scripts, and visualization of ORFs within contigs was performed in Geneious Basic [73] (Version 5.6.4). VF gene annotation, quantification, and visualization were implemented referencing the Virulence Factor Database (VFDB; Downloaded September 15, 2014) [59], following the same methods as ARG analysis.

Inferred interactions between phage and bacteria (**Script P11**, **Script R15**). A network of correlations between the relative abundances of bacterial genera from MetaPhlAn output and UniProt TrEMBL classified phages was constructed with CoNet [60] in Cytoscape v3.1.1 [61]. Bacterial and phage abundances were input into CoNet as separate matrices and taxa that were not present in at least 84 samples (based on recommended computed matrix information and specified with the minimum row filter) were filtered, and the sum of filtered rows was retained. We selected five different methods (Pearson and Spearman correlation metrics, the Mutual Information similarity metric, and Bray Curtis and Kullback-Leibler distance metrics) for ensemble inference of the network. Multiple measures were used to reduce false correlations and compositional biases. Thresholds were set automatically so that each method contributed the 250 top-ranking and 250 bottom-ranking edges to the network. P-values were computed from method- and edge-specific permutation and bootstrap score distributions, as follows. A random score distribution was generated using 100 permutations with the edgeScores routine, the row shuffling resampling method, and the renormalization option. The distribution was run with 100 bootstraps, p-values from the multiple metrics were combined using Simes' method [74], and FDR correction was performed [75]. Unstable edges, with edge scores outside of the 2.5 and 97.5 percentiles of the bootstrap distribution, were removed. Only interactions supported by two or more of the metrics specified above were retained. Network analysis was performed with the Cytoscape NetworkAnalyzer plugin [76].

CRISPR identification and comparison to the virome (**Script P12, Script R16**).

Putative CRISPR arrays were identified using PilerCR [62]. Consensus repeat sequences were extracted from the PilerCR output and exact duplicate sequences, reverse complements, and repeats less than 20nt long were removed. When repeats only differed by 2 nucleotides on either end, the shorter repeat was retained. Spacers ≤100 nucleotides long were identified by flanking repeats and extracted from the metagenome individual sample contigs. In order to identify viral targets, the spacers were queried against each viral contig using blastn. Because the spacer sequences are short, matches of 97% identity or greater were required, and hits deviating >3 nucleotides in length were rejected. Metagenomic reads containing CRISPR spacers were queried against the NCBI non-redundant database (blastn; $e<10^{-10}$) for host taxonomic classification. Metagenome-virome CRISPR interaction plots were generated using Circos [77]. To determine whether the CRISPRs were targeting coding or non-coding regions in phage, CRISPR spacers were queried against ORFs from the viral contigs that the spacers mapped to (blastn; $e<10^{-10}$, 97% identity). ORFs targeted by spacers were queried against the UniProt TrEMBL database using blastx ($e<10^{-10}$).

3.13 References

1. Hannigan GD, Grice EA: **Microbial ecology of the skin in the era of metagenomics and molecular microbiology.** *Cold Spring Harb Perspect Med* 2013, **3**.
2. Oh J, Byrd AL, Deming C, Conlan S, Program NCS, Kong HH, Segre JA, Program NCS: **Biogeography and individuality shape function in the human skin metagenome.** *Nature* 2014, **514:**59-64.
3. Bohannan BJ, Lenski RE: **Effect of Resource Enrichment on a Chemostat Community of Bacteria and Bacteriophage.** *Ecology* 1997, **78:**12.
4. Rodriguez-Brito B, Li L, Wegley L, Furlan M, Angly F, Breitbart M, Buchanan J, Desnues C, Dinsdale E, Edwards R, et al: **Viral and microbial community dynamics in four aquatic environments.** *ISME J* 2010, **4:**739-751.
5. Tyson GW, Banfield JF: **Rapidly evolving CRISPRs implicated in acquired resistance of microorganisms to viruses.** *Environ Microbiol* 2008, **10:**200-207.
6. Oliver KM, Degnan PH, Hunter MS, Moran NA: **Bacteriophages encode factors required for protection in a symbiotic mutualism.** *Science* 2009, **325:**992-994.
7. Tyler JS, Beeri K, Reynolds JL, Alteri CJ, Skinner KG, Friedman JH, Eaton KA, Friedman DI: **Prophage induction is enhanced and required for renal disease and lethality in an EHEC mouse model.** *PLoS Pathog* 2013, **9:**e1003236.
8. Brussow H, Canchaya C, Hardt WD: **Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion.** *Microbiol Mol Biol Rev* 2004, **68:**560-602, table of contents.
9. Hurwitz BL, Brum JR, Sullivan MB: **Depth-stratified functional and taxonomic niche specialization in the 'core' and 'flexible' Pacific Ocean Virome.** *ISME J* 2014.
10. Modi SR, Lee HH, Spina CS, Collins JJ: **Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome.** *Nature* 2013, **499:**219-222.
11. Foulongne V, Sauvage V, Hebert C, Dereure O, Cheval J, Gouilh MA, Pariente K, Segondy M, Burguiere A, Manuguerra JC, et al: **Human skin microbiota: high diversity of DNA viruses identified on the human skin by high throughput sequencing.** *PLoS One* 2012, **7:**e38499.
12. Wylie KM, Mihindukulasuriya KA, Zhou Y, Sodergren E, Storch GA, Weinstock GM: **Metagenomic analysis of double-stranded DNA viruses in healthy adults.** *BMC Biol* 2014, **12:**71.
13. Breitbart M: **Marine viruses: truth or dare.** *Ann Rev Mar Sci* 2012, **4:**425-448.
14. Thurber RV, Haynes M, Breitbart M, Wegley L, Rohwer F: **Laboratory procedures to generate viral metagenomes.** *Nat Protoc* 2009, **4:**470-483.
15. Reyes A, Haynes M, Hanson N, Angly FE, Heath AC, Rohwer F, Gordon JI: **Viruses in the faecal microbiota of monozygotic twins and their mothers.** *Nature* 2010, **466:**334-338.
16. Minot S, Sinha R, Chen J, Li H, Keilbaugh SA, Wu GD, Lewis JD, Bushman FD: **The human gut virome: inter-individual variation and dynamic response to diet.** *Genome Res* 2011, **21:**1616-1625.
17. Breitbart M, Haynes M, Kelley S, Angly F, Edwards RA, Felts B, Mahaffy JM, Mueller J, Nulton J, Rayhawk S, et al: **Viral diversity and dynamics in an infant gut.** *Res Microbiol* 2008, **159:**367-373.

18. Breitbart M, Hewson I, Felts B, Mahaffy JM, Nulton J, Salamon P, Rohwer F: **Metagenomic analyses of an uncultured viral community from human feces.** *J Bacteriol* 2003, **185:**6220-6223.
19. Roux S, Krupovic M, Debroas D, Forterre P, Enault F: **Assessment of viral community functional potential from viral metagenomes may be hampered by contamination with cellular sequences.** *Open Biology* 2013, **3**.
20. Minot S, Bryson A, Chehoud C, Wu GD, Lewis JD, Bushman FD: **Rapid evolution of the human gut virome.** *Proc Natl Acad Sci U S A* 2013, **110:**12450-12455.
21. Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, Huttenhower C: **Metagenomic microbial community profiling using unique clade-specific marker genes.** *Nat Methods* 2012, **9:**811-814.
22. Huson DH, Mitra S, Ruscheweyh HJ, Weber N, Schuster SC: **Integrative analysis of environmental sequences using MEGAN4.** *Genome Res* 2011, **21:**1552-1560.
23. Consortium HMP: **Structure, function and diversity of the healthy human microbiome.** *Nature* 2012, **486:**207-214.
24. Findley K, Oh J, Yang J, Conlan S, Deming C, Meyer JA, Schoenfeld D, Nomicos E, Park M, Becker J, et al: **Topographic diversity of fungal and bacterial communities in human skin.** *Nature* 2013.
25. Grice EA, Kong HH, Conlan S, Deming CB, Davis J, Young AC, Program NCS, Bouffard GG, Blakesley RW, Murray PR, et al: **Topographical and temporal diversity of the human skin microbiome.** *Science (New York, NY)* 2009, **324:**1190-1192.
26. Angly FE, Felts B, Breitbart M, Salamon P, Edwards RA, Carlson C, Chan AM, Haynes M, Kelley S, Liu H, et al: **The marine viromes of four oceanic regions.** *PLoS Biol* 2006, **4:**e368.
27. 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:**1694-1697.
28. Grice EA, Kong HH, Conlan S, Deming CB, Davis J, Young AC, Bouffard GG, Blakesley RW, Murray PR, Green ED, et al: **Topographical and temporal diversity of the human skin microbiome.** *Science* 2009, **324:**1190-1192.
29. Leplae R, Hebrant A, Wodak SJ, Toussaint A: **ACLAME: a CLAssification of Mobile genetic Elements.** *Nucleic Acids Res* 2004, **32:**D45-49.
30. Kanehisa M, Goto S: **KEGG: kyoto encyclopedia of genes and genomes.** *Nucleic Acids Res* 2000, **28:**27-30.
31. Abubucker S, Segata N, Goll J, Schubert AM, Izard J, Cantarel BL, Rodriguez-Mueller B, Zucker J, Thiagarajan M, Henrissat B, et al: **Metabolic reconstruction for metagenomic data and its application to the human microbiome.** *PLoS Comput Biol* 2012, **8:**e1002358.
32. Hurwitz BL, Brum JR, Sullivan MB: **Depth-stratified functional and taxonomic niche specialization in the 'core' and 'flexible' Pacific Ocean Virome.** *ISME J* 2015, **9:**472-484.
33. Schloss PD, Handelsman J: **A statistical toolbox for metagenomics: assessing functional diversity in microbial communities.** *BMC Bioinformatics* 2008, **9:**34.
34. Pride DT, Salzman J, Haynes M, Rohwer F, Davis-Long C, White RA, 3rd, Loomer P, Armitage GC, Relman DA: **Evidence of a robust resident bacteriophage population revealed through analysis of the human salivary virome.** *ISME J* 2012, **6:**915-926.

35. McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, Bhullar K, Canova MJ, De Pascale G, Ejim L, et al: **The comprehensive antibiotic resistance database.** *Antimicrob Agents Chemother* 2013, **57:**3348-3357.
36. Brotz E, Kulik A, Vikineswary S, Lim CT, Tan GY, Zinecker H, Imhoff JF, Paululat T, Fiedler HP: **Phenelfamycins G and H, new elfamycin-type antibiotics produced by Streptomyces albospinus Acta 3619.** *J Antibiot (Tokyo)* 2011, **64:**257-266.
37. Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, Jin Q: **VFDB: a reference database for bacterial virulence factors.** *Nucleic Acids Res* 2005, **33:**D325-328.
38. Soffer N, Zaneveld J, Vega Thurber R: **Phage-bacteria network analysis and its implication for the understanding of coral disease.** *Environ Microbiol* 2015, **17:**1203-1218.
39. Barabasi AL, Albert R: **Emergence of scaling in random networks.** *Science* 1999, **286:**509-512.
40. Zhou J, Deng Y, Luo F, He Z, Tu Q, Zhi X: **Functional molecular ecological networks.** *MBio* 2010, **1**.
41. Steele JA, Countway PD, Xia L, Vigil PD, Beman JM, Kim DY, Chow CE, Sachdeva R, Jones AC, Schwalbach MS, et al: **Marine bacterial, archaeal and protistan association networks reveal ecological linkages.** *ISME J* 2011, **5:**1414-1425.
42. Flores GE, Caporaso JG, Henley JB, Rideout JR, Domogala D, Chase J, Leff JW, Vazquez-Baeza Y, Gonzalez A, Knight R, et al: **Temporal variability is a personalized feature of the human microbiome.** *Genome Biol* 2014, **15:**531.
43. Gao Z, Tseng CH, Pei Z, Blaser MJ: **Molecular analysis of human forearm superficial skin bacterial biota.** *Proc Natl Acad Sci U S A* 2007, **104:**2927-2932.
44. Grice EA, Kong HH, Renaud G, Young AC, Bouffard GG, Blakesley RW, Wolfsberg TG, Turner ML, Segre JA: **A diversity profile of the human skin microbiota.** *Genome Res* 2008, **18:**1043-1050.
45. Kim KH, Bae JW: **Amplification methods bias metagenomic libraries of uncultured single-stranded and double-stranded DNA viruses.** *Appl Environ Microbiol* 2011, **77:**7663-7668.
46. Kim KH, Chang HW, Nam YD, Roh SW, Kim MS, Sung Y, Jeon CO, Oh HM, Bae JW: **Amplification of uncultured single-stranded DNA viruses from rice paddy soil.** *Appl Environ Microbiol* 2008, **74:**5975-5985.
47. Kim MS, Whon TW, Bae JW: **Comparative viral metagenomics of environmental samples from Korea.** *Genomics Inform* 2013, **11:**121-128.
48. Kleiner M, Hooper LV, Duerkop BA: **Evaluation of methods to purify virus-like particles for metagenomic sequencing of intestinal viromes.** *BMC Genomics* 2015, **16:**7.
49. Lood R, Collin M: **Characterization and genome sequencing of two Propionibacterium acnes phages displaying pseudolysogeny.** *BMC Genomics* 2011, **12:**198.
50. Marinelli LJ, Fitz-Gibbon S, Hayes C, Bowman C, Inkeles M, Loncaric A, Russell DA, Jacobs-Sera D, Cokus S, Pellegrini M, et al: **Propionibacterium acnes bacteriophages display limited genetic diversity and broad killing activity against bacterial skin isolates.** *MBio* 2012, **3**.
51. Willner D, Furlan M, Schmieder R, Grasis JA, Pride DT, Relman DA, Angly FE, McDole T, Mariella RP, Jr., Rohwer F, Haynes M: **Metagenomic detection of phage-encoded platelet-binding factors in the human oral cavity.** *Proc Natl Acad Sci U S A* 2011, **108 Suppl 1:**4547-4553.

52. Schmieder R, Edwards R: **Fast identification and removal of sequence contamination from genomic and metagenomic datasets.** *PLoS One* 2011, **6:**e17288.
53. Boisvert S, Raymond F, Godzaridis E, Laviolette F, Corbeil J: **Ray Meta: scalable de novo metagenome assembly and profiling.** *Genome Biol* 2012, **13:**R122.
54. UniProt C: **Activities at the Universal Protein Resource (UniProt).** *Nucleic Acids Res* 2014, **42:**D191-198.
55. Angly F, Rodriguez-Brito B, Bangor D, McNairnie P, Breitbart M, Salamon P, Felts B, Nulton J, Mahaffy J, Rohwer F: **PHACCS, an online tool for estimating the structure and diversity of uncultured viral communities using metagenomic information.** *BMC Bioinformatics* 2005, **6:**41.
56. **vegan: Community Ecology Package. R package version 2.2-0.** [http://CRAN.R-project.org/package=vegan]
57. Zheng Q, Wang XJ: **GOEAST: a web-based software toolkit for Gene Ontology enrichment analysis.** *Nucleic Acids Res* 2008, **36:**W358-363.
58. Delcher AL, Bratke KA, Powers EC, Salzberg SL: **Identifying bacterial genes and endosymbiont DNA with Glimmer.** *Bioinformatics* 2007, **23:**673-679.
59. Chen L, Xiong Z, Sun L, Yang J, Jin Q: **VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors.** *Nucleic Acids Res* 2012, **40:**D641-645.
60. Faust K, Sathirapongsasuti JF, Izard J, Segata N, Gevers D, Raes J, Huttenhower C: **Microbial co-occurrence relationships in the human microbiome.** *PLoS Comput Biol* 2012, **8:**e1002606.
61. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: **Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks.** *Genome Research* 2003, **13:**2498-2504.
62. Edgar R: **PILER-CR: Fast and accurate identification of CRISPR repeats.** *BMC Bioinformatics* 2007, **8:**18.
63. Hannigan GD, Hodkinson BP, McGinnis K, Tyldsley AS, Anari JB, Horan AD, Grice EA, Mehta S: **Culture-independent pilot study of microbiota colonizing open fractures and association with severity, mechanism, location, and complication from presentation to early outpatient follow-up.** *J Orthop Res* 2014, **32:**597-605.
64. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL: **BLAST+: architecture and applications.** *BMC Bioinformatics* 2009, **10:**421.
65. Huson DH, Mitra S, Ruscheweyh H-J, Weber N, Schuster SC: **Integrative analysis of environmental sequences using MEGAN4.** *Genome Research* 2011, **21:**1552-1560.
66. Langmead B, Salzberg SL: **Fast gapped-read alignment with Bowtie 2.** *Nat Methods* 2012, **9:**357-359.
67. Angly FE, Willner D, Prieto-Davo A, Edwards RA, Schmieder R, Vega-Thurber R, Antonopoulos DA, Barott K, Cottrell MT, Desnues C, et al: **The GAAS metagenomic tool and its estimations of viral and microbial average genome size in four major biomes.** *PLoS Comput Biol* 2009, **5:**e1000593.
68. Segata N, Boernigen D, Tickle TL, Morgan XC, Garrett WS, Huttenhower C: **Computational meta'omics for microbial community studies.** *Mol Syst Biol* 2013, **9:**666.
69. Wickham H: *ggplot2: elegant graphics for data analysis.* New York: Springer; 2009.
70. McGill R, Tukey J, Larsen W: **Variations of Box Plots.** *The American Statistician* 1978, **32:**12-16.

71. Edgar RC: **Search and clustering orders of magnitude faster than BLAST.** *Bioinformatics* 2010, **26:**2460-2461.
72. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, et al: **QIIME allows analysis of high-throughput community sequencing data.** *Nat Methods* 2010, **7:**335-336.
73. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al: **Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data.** *Bioinformatics* 2012, **28:**1647-1649.
74. Sarkar SK, Chang CK: **The Simes method for multiple hypothesis testing with positively dependent test statistics.** *Journal of the American Statistical Association* 1997, **92:**1601-1608.
75. 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* 1995, **57:**289-300.
76. Assenov Y, Ramirez F, Schelhorn SE, Lengauer T, Albrecht M: **Computing topological parameters of biological networks.** *Bioinformatics* 2008, **24:**282-284.
77. Krzywinski MI, Schein JE, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA: **Circos: An information aesthetic for comparative genomics.** *Genome Research* 2009.

CHAPTER 4 – An Introduction to Current Concepts and Ongoing Research in the Prevention and Treatment of Open Fracture Infections

The contents of this chapter have been published as:

> Hannigan GD, Pulos N, Grice EA, Mehta S. 2014. Current concepts and ongoing research in the prevention and treatment of open fracture infections. Advances in Wound Care. doi: 10.1089/ wound.2014.0531.

4.1 Abstract

4.1.1 Significance

Open fractures are fractures in which the bone has violated the skin and soft tissue. Because of their severity, open fractures are associated with complications that can result in increased lengths of hospital stays, multiple operative interventions, and even amputation. One of the factors thought to influence the extent of these complications is exposure and contamination of the open fracture with environmental microorganisms, potentially those that are pathogenic in nature.

4.1.2 Recent Advances

Current open fracture care aims to prevent infection by wound classification, prophylactic antibiotic administration, debridement and irrigation, and stable fracture fixation.

4.1.3 Critical Issues

Despite these established treatment paradigms, infections and infection-related complications remain a significant clinical burden. To address this, improvements need to be made in our ability to detect bacterial infections, effectively remove wound contamination, eradicate infections, and treat and prevent biofilm formation associated with fracture fixation hardware.

4.1.4 Future Directions

Current research is addressing these critical issues. While culture methods are of limited value, culture-independent molecular techniques are being developed to provide informative detection

of bacterial contamination and infection. Other advanced contamination- and infection-detecting techniques are also being investigated. New hardware-coating methods are being developed to minimize the risk of biofilm formation in wounds, and immune stimulation techniques are being developed to prevent open fracture infections.

4.2 Scope and Significance

Open fractures occur when bone is exposed through skin as a result of bone breaking through skin or wound penetration with fractured bone exposure. While multiple factors may influence open fracture rate, a recent study reported an incidence of 30.7/105/year [1]. Open fractures have multiple causes, often occur in extremities, and are most severe in lower legs and feet (Fig. 1A–C). Infection rates also vary, but have been reported as 2.3% with effective antibiotic treatment [2]. Because of their severity, open fractures are associated with complications, including longer hospital stays, multiple operative interventions, and amputations (average amputation lifetime healthcare costs over $500,000.) [3].

4.3 Translational Relevance

Microbes are known to complicate open fracture healing through infections and biofilm formation, as well as potentially playing roles in nonunion/malunion cases [4]. Further, while the importance of microbes in open fracture healing is accepted as significant, there is still little known about how or what microbes affect these wounds, or how we can use microbes diagnostically to predict complications and better inform treatment. The goal of current research is to address the deficiencies in the current paradigms of open fracture care, and to improve prevention and treatment of open fracture infection.

4.4 Clinical Relevance

Microbial contamination and infection are common concerns in all wound care scenarios, but open fractures are at a higher risk for infection and other microbe-related complications. Open fractures are often the result of high-energy events that result in severe bone and soft tissue damage, thereby leading to significant risk of infection. Open fracture care focuses on effective management, especially in the early stages, with the goal of minimizing complications caused by microbial contamination.

4.5 Discussion

4.5.1 Current concepts in open fracture care and infectious risk minimization

Open fractures are at a high risk for infection and other complications, and the steps taken during initial treatment have a significant impact on the overall outcome. This impact has been evidenced as a decrease in infections and other complications as the result of effective fracture classification, prophylactic antibiotic administration, early debridement and irrigation, and proper fracture fixation. In this section we will further discuss the importance of initial wound management, highlighting the current concepts in open fracture care and the standard treatments, both prophylactic and therapeutic, for infections during open fracture healing.

4.5.1.1 Open fracture classification and diagnosis

The initial description and evaluation of the wound is important for informing downstream actions and standardizing descriptive measures among the medical professional community. Several classification methods have been proposed for the description and evaluation of open fractures. The most frequently quoted and widely used scheme was first described by Gustilo and Anderson [5], and later modified to its current form by Gustilo et al. (Table 1) [6]. This classification system involves the intraoperative scoring of open fractures from one to three in ascending order of severity, with a Type I injury involving a small soft tissue wound, Type II

involving a large wound with little soft tissue damage, and Type III involving extensive soft tissue damage. Examples of Type II and Type III severity are shown in Fig. 2. Type III wounds are further subcategorized into three subgroups, with Type IIIa fractures having extensive soft tissue damage with adequate soft tissue coverage, Type IIIb fractures having extensive soft tissue damage requiring transfer of soft tissue to cover the defect, and Type IIIc being the most severe due to extensive arterial damage requiring vascular repair. The severity of the open fracture, as scored by the Gustilo and Anderson classification system, is associated with the rate of infection and therefore has prognostic value [5-7].

Oestern and Tscherne proposed a classification system based on fracture type and soft tissue damage for both open and closed fractures (Table 2) [8]. Additionally, the Association for the Study of Internal Fixation (translated from the German "Arbeitsgemeinschaft für Osteosynthesefragen" and abbreviated as the AO Foundation) has published a classification system that is designed to provide information about both the soft tissue and bone damage of the open fracture (Table 3) [9, 10]. This scheme considers the damage done to skin, the muscle tissues and tendons, and neurovascular system, overall making this a comprehensive and accurate classification scheme [9, 10]. While both classification methods are valuable, the Gustilo and Anderson classification scheme remains the most widely used due to its simplicity and familiarity.

4.5.1.2 Prophylactic antibiotic administration

Once an open fracture has been identified and loosely classified in the resuscitation bay or emergency room (typing of open fractures is most accurate in the operating room), treatment with antibiotics is initiated to minimize the risk for infection. Obvious gross contamination is also

removed at this time. These treatments are performed as early as possible after the traumatic event in order to minimize infection and other complications.

Infections and bacterially related complications are important concerns when treating open fractures. In fact, the Gustilo and Anderson study of 1976 reported positive initial bacterial cultures in 70.3% of the 158 prospectively observed open fractures.5 Because bacterial colonization was strongly associated with open fractures, orthopedic professionals accepted, without evidence, that prophylactic antibiotics would lower the risk of wound infection. Attempts to address the utility of prophylactic antibiotics yielded weak and conflicting results until 1974, when Patzakis et al. reported a reduction in open fracture wound infections from 13.9% in patients without antibiotic treatment, to 2.3% when patients were treated with cephalothin antibiotics [2]. This study strongly supported the need for prophylactic antibiotic use. The study also illustrated the importance of understanding the administered antibiotics because, while cephalothin antibiotics were significantly effective in reducing infection rate, the patients treated with penicillin and streptomycin did not show a significant reduction in infection rate (a nonsignificant reduction from 13.9% in nontreated patients to 9.7% in penicillin and streptomycin treated) [2]. Patzakis et al. used antibiotic resistance culture techniques to show this was at least partially due to penicillin and streptomycin resistance. Other studies have since highlighted other important considerations when deciding an antibiotic regimen.

Deciding an appropriate antibiotic course requires an understanding of the bacteria most likely to colonize wounds. Both Patzakis et al. and Gustilo and Anderson found that, when culturing wound infections, staphylococci (specifically coagulase positive staphylococci such as Staphylococcus aureus) were the most commonly isolated organisms [2, 5]. Because these bacteria appeared to be the most likely causes of infections, they suggested that prophylactic

antibiotics should target Gram positive bacteria, and most especially staphylococci. The benefits of prophylactic antibiotic use against Gram positive bacteria have since been supported by other series [11-13]. While research supports the benefits of prophylactic antibiotics that target Gram positive bacteria, there is insufficient evidence to support the prophylactic use of Gram negative antibiotics [9, 13, 14]. As Gram negative bacteria become more prevalent in open fracture infections, including Acinetobacter baumannii and Pseudomonas aeruginosa, research addressing the prophylactic benefits of Gram negative antibiotics will become increasingly important.

Another concern is the increasing threat of acquisition of antibiotic resistance by bacteria. The continued emergence of methicillin-resistant S. aureus (MRSA) has brought new considerations to prophylactic antibiotic treatment of open fractures [14]. Various rates of MRSA colonization of the nares have been reported, with a high rate of 7.4% in healthy university students in 2009 [15], and general rates being around 1–2.5%.16–20 MRSA colonization in the nares, axilla, and groin has been suggested to increase the risk of MRSA infection at surgical wound sites [14, 16], and nasal decolonization treatments, paired with antibiotic prophylaxis, have been shown to reduce the risk of MRSA infection in some cases [17]. Although the benefits of prophylactic antibiotic regimens that target MRSA have not yet been established, this will likely continue to be an important consideration as surgeons decide the best prophylactic antibiotic regimens to administer [14].

In addition to MRSA, there is also an increasing concern about other antibiotic-resistant bacteria, including Acinetobacter, Klebsiella, Pseudomonas, and Enterobacter, which are present in open fractures[18] and are known to be potentially infectious agents of open wounds [19-21]. Notably, pan-resistant strains of the significant hospital pathogen A. baumannii have emerged, developing resistance to colistin, the drug of last resort [22]. Antibiotic-susceptible and -resistant A.

baumannii infections have continued to increase in prevalence over the past decades, both in military and civilian settings [23-25]. Because the antibiotic-resistant profiles of A. baumannii and other potentially antibiotic-resistant bacteria can vary geographically, orthopedic clinicians must consider the local potentially pathogenic bacteria and the local antibiotic-resistance profiles associated with those bacteria, as has been suggested for treating open fracture A. baumannii infections [20].

While it is important to predict what prophylactic antibiotics will be most effective, it is also important to understand the ideal administration timeline. This timeline includes the ideal gap length between patient presentation and antibiotic administration, and duration of antibiotic administration. Most surgeons agree that prophylactic antibiotic treatment should be started as soon as possible [7, 9, 14]. The key study by Patzakis and Wilkins showed that the most important treatment in preventing open fracture infection is prophylactic antibiotic administration [7]. The group showed that the patients who were treated prophylactically with antibiotics within 3 h were less likely to develop infection, and this timeframe is still used today.

While ideal time to administration is straightforward, ideal duration of therapy is less clear. One study has suggested that antibiotic treatment should be continued for 3 days after initiation [26], while another study argues that 24 h is no less effective than 72 h [27]. Currently, authors advise that antibiotic treatment should be continued for at least 24 h, and may be continued for up to 72 h [9, 13, 14]. A concern with longer antibiotic administration times (i.e., 72 h) is that the increased exposure may promote antibiotic resistance among the bacterial populations, which has been shown to occur in some cases [9, 13]. Further research will be required to provide definitive responses to these concerns.

4.5.1.3 Debridement and irrigation

In addition to prophylactic antibiotic administration, wound debridement and irrigation are important procedures for preventing open fracture infection. The goal of open fracture surgical debridement is the excision of environmental debris, devitalized soft tissue, and bone, as well as irrigation of the wound to reduce bacterial load. The three major considerations are ideal timing between injury and debridement, the extent of debridement, and the irrigation materials to be used.

The recommended time to debridement after injury is dictated by the "6-h rule." The 6-h rule is an orthopedic rule of thumb that claims that, to be effective, open fracture debridement should be conducted within 6 h after the injury. While this rule is widely accepted, little scientific evidence supports it. The 6-h rule was started by Friedrich, who utilized a guinea pig model and reported decreased risk of infection when contaminated soft tissue wounds were debrided <6 h after contamination. This suggested that debridement earlier than 6 h resulted in lower infectious risk [28]. Recent literature, including human studies, has not supported the Friedrich claims, and while further study is needed, it seems that there is no increased risk of infection in delayed debridement cases [9, 29]. Despite these recent findings, most surgeons recommend immediate debridement of highly contaminated types II and III open fractures [9].

Irrigation is an important supplement for aggressive debridement of necrotic tissue and particulate matter because it further removes particulate debris and bacteria from the wound. The specifics of what materials should be used, and to what extent irrigation should be performed, remain a topic of debate. These discussions focus on whether soaps, antiseptics, or antibiotics should be included in the irrigation saline, and whether pulsatile lavage should be used. Pulsatile lavage is a point of

concern because it may drive bacteria further into soft tissue and cause microscopic damage to the soft tissue, and thereby impede healing and increase the risk for infection.

Most surgeons irrigate open fracture wounds using sterile saline alone, saline in combination with soap, saline in combination with antiseptic chemicals, or saline in combination with antibiotics. Although one might intuitively think that chemical additives would eliminate more bacteria and decrease the risk of infection, they are found to be ineffective. Antiseptic compounds are known to destroy bacteria, but studies have yielded conflicting results regarding their beneficial effect on wound healing and infection rate compared to saline solutions alone [30, 31]. In fact, antiseptics may be toxic to the human host cells, which could limit their efficacy due to host damage [30, 31]. Like antiseptics included in saline, the beneficial effects of human wound irrigation using antibiotics with saline has been unconvincing [30, 32].

Unlike antiseptics and antibiotics that destroy bacteria, soaps facilitate the physical removal of bacteria. Studies suggest that the use of soap with saline is just as, if not more, effective compared to antibiotic saline solutions [32-34]. Because antibiotic and antiseptic use in saline can add additional cost to treatment, may promote antibiotic resistance, and may harm the human host tissue, orthopedic clinicians recommend the use of soap with saline when irrigating wounds [9].

The pressure used in open fracture irrigation is just as important as the solutions used. The benefits of low-pressure gravity irrigation or high-pressure pulsatile lavage in open fractures remain a point of discussion. While high-pressure pulsatile lavage seems attractive because it is thought to better remove entrenched bacteria and debris, the high pressure may push bacteria further into the tissue. High-pressure irrigation may also heighten the risk of healing complications because it damages the surrounding human tissue. Although bacteria and debris

may be more effectively removed from wounds using high-pressure pulsatile lavage [35], others have argued that high-pressure pulsatile lavage does in fact push bacteria further into tissue and increases the numbers of bacteria retained in the wound [36]. Additionally, there has been significant research to suggest that high-pressure pulsatile lavage damages human tissue, thereby increasing risks for complications, infections, and delayed healing [36, 37]. Together, the effects of high-pressure irrigation are seen as more destructive than helpful, and low-pressure irrigation is recommended [9].

Volume of irrigation solution used is also an important consideration. In 2001, Anglen proposed increasing volumes of irrigation for more severe wounds. Given the availability of 3-L irrigation bags, he proposed 3 L for Type I fractures, 6 L for Type II fractures, and 9 L for Type III fractures [30]. This is the method currently used by most surgeons.

4.5.1.4 Internal and external fracture fixation devices

After the initial treatment of the wound, attention turns to fracture reduction and fixation. Not only is fracture reduction (anatomical realignment of fracture fragments) important for proper bone union and healing, stabilization of the fractured bones limits soft tissue damage. Anatomic reduction mediates the inflammatory response, decreases hematoma volume and dead space, and allows for tissue revascularization.

The utilization and efficacy of various fracture fixation techniques differs based on the anatomical location and severity of the injury. There are three general methods to fixing fractures: plate fixation, intramedullary (IM) nailing, and external fixation. Examples of IM nailing and external fixation are shown in Fig. 3. Both plate fixation and IM nailing are internal fixation approaches, while external fixation is external, as the name suggests. Due to the high rates of

complications and concerns for periosteal blood supply damage associated with plate fixation methods, they have been largely replaced by IM nailing and external fixation techniques for lower extremity diaphyseal fractures [38]. However, plates are still commonly used for periarticular fractures and open fractures of the radius and ulna, as it becomes more important to get an anatomic reduction. For example, internal plate fixation has been shown to be more effective than external fixation in the treatment of distal radius fractures [39].

The external fixation technique involves the insertion of threaded pins into the fractured long bones through the skin (Fig. 3B). These screws are attached to external hardware that provides stable fracture fixation. The advantages of the external fixation approach are that it allows for rapid fracture stabilization, avoids placement of internal hardware, and minimizes further soft tissue damage by placing screws outside the zone of injury. Pin-track infections, concerns about fracture malalignment, and poor patient compliance limit its use for definitive fixation. External fixation is now more commonly used for temporary fixation of fractures while the surgeon awaits the soft tissues to recover, eventually converting to internal fixation.

IM nailing is an internal fixation approach for long-bone fractures, in which a titanium or stainless steel rod is placed into the reamed or unreamed medullary canal of a long bone (Fig. 3A). This rod is secured in place and serves as an internal scaffold around which bone can heal. The advantage to the IM nailing technique is that it offers effective bone fixation that maintains length, alignment, and rotation, and also allows for earlier weight bearing. Though reamed femoral nailing is the gold standard for closed femoral shaft fractures [40], concerns about infection risk in open fractures have been raised [41]. However, two prospective randomized trials do not show a significant increase in infection risk when using a reamed, locked IM nail for treatment of open tibial shaft fractures [42, 43]. With different technical advantages to each

fracture fixation technique, the surgeon must take into account fracture pattern and soft tissue injury when deciding which method will best provide a positive functional outcome [38, 44].

4.5.1.4 Treatment of fracture and soft tissue infections

Following initial classification, prophylactic antibiotic administration, surgical debridement and irrigation, and fracture stabilization, the open fracture wound may still become infected. In this case the patient is most often treated with intravenous antibiotics to suppress and eliminate the infection. Depending on the nature of the infection (the severity, location, and depth), the fracture fixation hardware may be left in place until the fracture heals, and will be removed after healing. If the infection is more severe, then the hardware may have to be removed, the wound will be debrided in addition to local antibiotic administration, and the hardware will be reinstalled after the infection has been cleared. Another common infectious concern is the formation of biofilms, which can occur rapidly on medical devices as well as host substrates like bone.

4.5.2 Ongoing research and the future of open fracture care

Although there are many established treatment paradigms in place for open fractures, infections and other complications remain a present threat. Research to improve these treatments remains ongoing. Now that we have discussed the more established concepts in current open fracture care, we are going to move our focus toward the ongoing dilemmas facing open fractures and infection prevention/treatment, and the research aimed at finding solutions.

4.5.2.1 Detection of bacterial contamination at time of injury

One of most comprehensive problems in management and treatment of open fractures is identification and quantification of microbial contamination at time of injury. Identifying microbial biomarkers indicative of complication risk would also better inform open fracture

management. Surveillance cultures at the time of presentation have little value in predicting what organism will cause a downstream infection. One study that illustrated this limited predictive value of surveillance cultures was a prospective clinical study by Valenziano et al [45]. The group collected swabs from open fractures upon patient presentation to the hospital (before antibiotic intervention), obtained aerobic and anaerobic cultures from the samples, and examined correlations between the cultures and the patient progressions to infection. Only 24% of the surveillance cultures resulted in growth. Additionally, 77% of the infected wounds yielded negative cultures, and none of the cultured organisms matched the infectious organisms. This suggested an inability of surveillance cultures to reliably predict the infectious organisms of open fractures. This inability of surveillance cultures to accurately predict the infectious organism has been supported by other studies [45-48].

While surveillance cultures have limited value in predicting downstream infectious organisms, some studies have suggested a value in surveillance culture bacterial load quantification. This was recently addressed in a retrospective study conducted by Burns et al. in a combat environment. The group took a similar approach to that mentioned previously, by attempting to find correlations between surveillance cultures taken during initial wound debridement and the later development of infections. Burns et al. found that the positive surveillance cultures were not able to accurately predict the infecting organism, as has been shown before. However, 38.7% of the culture-positive patients went on to infection, while only 11.5% of the culture-negative patients developed infection, and this correlation between a positive bacterial load culture and progression to infection was found to be significant. This therefore suggested that quantitative bacterial culturing may have limited value in predicting general infection. Other studies have also demonstrated the value of quantifying bacterial loads of wounds for general infection prediction, either through the use of quantitative Gram staining or more commonly through the use of

quantitative culturing [48-51]. The samples used for these bacteria quantifications were either wound swabs, wound effluent, or debrided tissue.

The timing of sample collection for quantification, such as whether the sample is collected before or after debridement, may be important and this may explain some different results reported in the literature. In a study by Merritt, the surveillance cultures for bacterial load were shown to have predictive value when taken as the patient was leaving the operating room (after the wound was debrided, irrigated, and cleaned), but not when taken as the patient entered the operating room (the sample was taken during debridement) [51]. This suggested that the timing of surveillance culture sampling may be important. Although this study was conducted many years ago, the importance of sampling timing will likely remain a point for further investigation.

4.5.2.2 Advances in molecular analysis of bacterial contamination at time of injury

The advent of next-generation sequencing platforms, with increased throughput and decreased costs, has enabled approaches that do not rely on cultures for bacterial identification. Based on the DNA sequence of the prokaryote-specific 16S small subunit ribosomal RNA (rRNA) gene, culture-independent sequencing methods eliminate biases associated with cultures. Our group recently reported an ongoing pilot study that is using such approaches to understand the bacteria associated with open fractures [52]. This study utilized high-throughput sequencing of the bacterial 16S rRNA gene to characterize 30 open fractures, and was able to correlate specific bacterial taxa and community dynamics with time points and other clinical factors, including the anatomical wound location and patient progression to healing complications. The data can also be used to visualize the differences in bacterial communities between anatomical sites, and between the wound and healthy skin, at presentation of the patient to the emergency room (Fig. 4). It also shows that healthy skin communities are dominated by Corynebacteriaceae and

Staphylococcaceae bacteria, while the wound communities are not strongly dominated by any particular bacteria (Fig. 4). Because this was a pilot study, the prognostic value of certain bacterial abundances or community compositions were not addressed, but this will be an obvious next step as more patients are enrolled and as more follow-up information is collected until the end of each patient's healing process. Overall, this study is allowing for more robust, detailed studies of the communities associated with open fractures.

Just as it is important to understand the specific bacteria that contaminate open fractures and cause infectious complications, it is also important to understand the ecology of open fracture wound bacterial communities. Up to this point, individual-cultured bacteria have been primarily considered either harmful or potentially pathogenic. In fact, not all bacteria are harmful, and some can be beneficial. Having a better understanding of these harmful and beneficial groups will improve further therapeutic development.

Changes in the human microbiome have been associated with a multitude of inflammatory diseases and states, including inflammatory bowel disease, acne vulgaris, and atopic dermatitis [53, 54]. In these cases, disease states are associated with alterations in the bacterial community structure, an alteration referred to as “dysbiosis.” Together, these examples illustrate that the entire microbial community, not just the potentially pathogenic or opportunistic microorganisms, influences host–microbe homeostasis. Further, commensal bacteria are thought to promote health in many ways, including competitive inhibition of potentially opportunistic and/or pathogenic microorganisms, educating and modulating the host immune response, and through the production of compounds that inhibit growth of potential pathogens, such as antimicrobial peptides. Work toward understanding the open fracture microbiome, and the beneficial and harmful bacteria in that community, remains ongoing.

Perhaps one of the most basic culture-independent, molecular methods is the estimation of bacterial load using quantitative polymerase chain reaction techniques. This method involves the quantification of the 16S rRNA gene sequences present in a wound swab or other sample type. This method is also ideal as a basic starting point because it does not require any sequencing of the bacterial genome, as primers are designed to regions of the 16S rRNA gene that are conserved throughout a broad range of prokaryotic taxa. Although this method is used in research laboratories [4, 52], it has not yet been implemented in clinical settings.

Another high-throughput approach to understanding host–microbe homeostasis in traumatic injury was reported in a recent study by Chromy et al., who investigated the utility of global protein profiling approaches for identifying host biomarkers [55]. Wound effluent was collected prior to, and shortly after, surgical debridement. Twenty-five proteins were significantly differentially expressed between uneventful healing and complicated healing groups, many with established roles in regulating inflammatory and immune responses. For example, increased expression of complement C3 protein was associated with dehisced wounds, a similar finding to a chronic wound model in which complement genes were upregulated [56]. Excessive complement activation can be damaging to the host and has been linked to myriad inflammatory and autoimmune conditions [57]. Although these identified host biomarkers need further validation in open fracture settings (only one open fracture was included in the study of 19 patients with severe traumatic injury), this general approach is promising as a readout of the host immune response and may enable identification of protein biomarkers with predictive and/or prognostic value.

4.5.2.3 Techniques and methods for contamination eradication

While it will be important to continue to improve methods for diagnosing contaminated wounds and predicting their outcomes, it will also be important to improve methods for eradicating contamination from open fracture wounds. Open fracture infection is currently prevented through the minimization of contamination, often practiced as aggressive wound debridement, irrigation, and prophylactic antibiotic administration. The use of local antibiotic therapy for severe open fractures (Type IIIB and Type IIIC) has been shown to reduce the incidence of infection in a series of 1,085 open fractures.62 Ostermann et al. used aminoglycoside-impregnated polymethylmethacrylate (PMMA) beads to provide high local concentrations of antibiotics. Because PMMA is not bioabsorbable, the length of implantation remains controversial and requires retrieval. Bioabsorbable antibiotic delivery vehicles may eliminate the need for reoperation and removal [58].

4.5.2.4 Assessment of open fractures for infection

Just as it is important to accurately diagnose bacterial contamination at the time of injury, it is also important to accurately assess wounds for infection. Although it may seem the assessment of infection should be obvious, this remains a difficult procedure. In fact, a series of studies that began in 1995 showed that, in cases of otitis media, inflammatory and bacterial cells could be observed by microscopy, the presence of bacteria could be confirmed by 16S rRNA gene quantification, and the presence of live bacteria could be confirmed by mRNA quantification, but the majority of bacterial cultures remained negative [59]. Due to these inaccuracies, culture methods alone are not sufficient to properly diagnose an infection. In fact, no single method is sufficient for infection diagnosis and multiple methods must be used for proper diagnosis [59-61]. Methods for the assessment of infections include repeated measurements of immune-related markers (i.e., C-reactive protein and erythrocyte sedimentation rate), culturing, histopathology, X-ray imaging (diffuse periosteal reaction, fracture delayed union or nonunion, or loosening of

pins indicates potential infection), nuclear imaging of 99mTc accumulation, and computed tomography, including magnetic resonance imaging and positron emission tomography methods [59-61]. Because of the level of specialty required, infectious disease teams will often coordinate with the orthopedic team, when available, to identify and provide the most appropriate treatment. An overview of the mentioned methods for detecting open fracture bacterial contamination and infection can be found in Table 4.

4.5.2.5 Prevention of biofilm formation on hardware

A specific infectious interest to orthopedic clinicians is the prevention of biofilms on fracture fixation hardware. Biofilms are complex communities of bacteria that create extracellular polymers that allow them to adhere to each other, as well as to implanted devices. Biofilms are a particular concern in open fractures, as well as other implant settings, because they are difficult to eradicate. Most antibiotics are unable to penetrate into biofilms, thereby weakening the primary line of attack. Biofilms also make the enclosed bacteria resistant to most effects of the host immune system. Additionally, the close proximity of biofilms creates an environment that promotes horizontal gene transfer, including transfer of antibiotic resistance and other virulence factor genes [62]. Culture identification of microorganisms forming a biofilm is challenging, as those microbes forming the biofilm rely on microbe–microbe interactions, and are thus difficult to isolate as individual planktonic colonies. Additionally, biofilms are usually polymicrobial, and are often collections of Gram positive and negative bacteria, which makes their culture identification and treatment particularly difficult.

Biofilms are polymicrobial and maintain a “supragenome” that is necessary for the overall biofilm survivability [63]. This means that biofilms are complex communities of bacteria that, together, express the genes needed for biofilm formation and maintenance, but no single bacteria

has all of the required genes; the genetic burden is shared among the community [63]. Because of the metagenomic synergy, bacterial diversity, horizontal gene transfer, and overall genomic diversity associated with biofilms, almost any bacteria is capable of forming a biofilm. All of these factors contribute to the difficulty in treating biofilm infections. Overall, the best approach is to prevent biofilm formation in the first place.

Open fractures are at a higher risk for biofilm infections compared with closed fractures, likely because they have a greater burden of contamination and deficient immune responses. Incidence of biofilm formation after open fracture internal fixation may exceed 30% [60]. Prevention of biofilm formation is important because biofilms can delay healing, propagate complications, and increase treatment costs. While there are multiple methods to prevent biofilm formation, including prophylactic antibiotics and accurate detection of potential biofilm-forming bacteria (both discussed previously), we will focus on hardware coatings that can deter or prevent bacterial adhesion and biofilm formation.

Biofilm-prevention studies are conducted on many different types of devices, but recently the group of Williams et al. reported an effective antimicrobial coating that was tested in a type IIIB open fracture sheep model [64]. The coating reported in this study was an active release compound (meaning the coating continuously releases the antimicrobial compounds into the surrounding tissue) that was composed of silicone polymer and an active release antimicrobial agent called cationic steroid antimicrobial-13. Williams et al. found that their coated fracture fixation devices prevented 100% of infections when challenged with biofilm inocula in the open fracture sheep model, and 100% of the uncoated devices went on to infection. This particular coat shows promise and warrants further investigation.

Other coat-based approaches to preventing biofilm formation include the use of antisense molecules that can target and silence bacterial virulence factor genes, the use of quorum sensing inhibitors, and even coating with bacteriophages (viruses that only target and destroy bacteria), which are capable of penetrating biofilms [62]. Additionally, the use of ultrasound or electric currents may be effective in disrupting biofilms to allow for antibiotic or antimicrobial compound penetrance [62]. The use of external fixation devices as discussed previously, when possible, is another way to reduce the risk of biofilm formation. External fixation devices can potentially reduce the risk for infection because the pins are placed outside of the zone of injury, because they have a smaller surface area, and because they are never permanent. Because pin-site infections are common, these devices' values are also limited.

4.5.2.6 Difficulties in elimination of infection and biofilm destruction without sacrificing construct stability

When biofilm infections do occur on internal fracture fixation hardware, the treatment must balance the risks of fostering infection with the benefits of fracture stability. As mentioned earlier, treatment of biofilms is particularly difficult because the structure protects the bacteria from antibiotics and host immune responses. The choice to remove hardware to treat a potential hardware biofilm infection depends largely on the state of bone healing. If the patient's bone has sufficiently healed, then the hardware is removed and the patient is treated with antibiotics. The case becomes more difficult when the bone has not sufficiently healed.

If the patient's bone has failed to heal, then the surgeon must make a decision as to whether the hardware should remain until the bone has healed, or to remove the hardware, treat the infection (often with local antibiotics actively released by a PMMA vehicle), and install new hardware to

stabilize the fracture after the infection has cleared. This can be a difficult choice and in many cases an infectious disease specialist is consulted.

4.5.2.7 Increased risks of infection due to deficient immune responses in open fractures

One of the major ways open fracture wounds are left more susceptible to infection is their deficient immune response. While there are multiple deficiencies in the local immune responses after an open fracture, one of the deficiencies is decreased function of T helper 1 (TH1) lymphocytes. TH1 lymphocytes are important modulators of the cellular immune response, as well as the production of complement-fixing antibodies [65]. This deficiency in TH1 lymphocytes has been linked to the reduced ability of open fractures to resist infections, and attempts to restore TH1 function in open fractures have resulted in increased resistance to infection in animal models [65, 66]. This knowledge has led some groups to attempt to prevent open fracture infection by modulating the immune system.

In 2012, a group led by Boyce et al. attempted to therapeutically modulate the immune response in an open fracture rat model, in which the rats' femurs were fractured using a custom apparatus [65]. The group used IL-12 to modulate the immune response because IL-12 is known to play a role in naive T lymphocyte differentiation into TH1 lymphocytes, which would therefore stimulate the wounds' immune response and address their deficiency in TH1 lymphocytes. After femur fracture, the group inoculated the wounds with clinical isolates of S. aureus and treated the rat wounds with percutaneous injections of placebo, IL-12, ampicillin antibiotic, or a combination of IL-12 and antibiotic. The group found that, although the antibiotic treatment was more effective than IL-12 alone in preventing infection, the combination of IL-12 and antibiotic was more effective than the antibiotic treatment alone. This suggests that the use of IL-12, in combination with antibiotic treatment of open fractures, may improve the wound's resistance to

infection. While this was only an animal model study, it warrants further investigation into using immune-modulating cytokines to improve the efficiency of prophylactic antibiotic, or other antibiotic treatments.

The same group, led by Li et al., also investigated the efficacy of coating implant devices with IL-12 to prevent biofilm formation and infection by stimulating the immune system as described previously [66]. The group used the rat femur fracture model and S. aureus bacterial challenge model as described previously. Metallic wires were used as IM nails at the fracture sites. Half of the rats received wires coated with IL-12 and the other half received uncoated wires. The results showed that rats who received IL-12-coated wires had significantly lower rates of infection, and those rats also had better bone quality and improved healing as assessed by three blinded, orthopedic surgeons. This report supports the benefits of IL-12 as coatings on implant devices, such as IM nails, and warrants further investigation.

4.5.3 Conclusions and perspectives

Due to exposure to the external environment, the extended duration of required healing, and suppressed immune responses, open fractures are at significant risk for infectious complications. A major focus of current open fracture care is minimization of this infectious risk. During initial treatment, infectious risk of the open fracture is reduced by properly categorizing the wound, treating the patient with prophylactic antibiotics, debriding and irrigating the wound, and stabilizing the fracture with appropriate hardware.

The early detection of bacterial contaminants continues to be a focus of current research. Unfortunately, contemporary surveillance culture methods are unable to reliably predict the bacteria that will lead to infection, often because the cultured bacteria are not the same bacteria

present at the time of infection. There is still a need to accurately predict which patients will move on to develop infections of particular bacteria, and researchers will likely continue to investigate potential methods for making such predictions.

Timing of sample collection will likely play a role in the success of biomarker discovery for infecting bacteria. Timing of sample collection, such as whether the sample was taken before or after surgical debridement, is a potentially significant factor in whether or not the detected bacteria will lead to downstream infections. Further, existing studies have focused on detecting potentially infectious organisms upon presentation, or shortly thereafter, but often fail to assess the potentially infectious organisms colonizing the wound at later times. This may be important because the bacteria present at the wound site at later times may be more significant to causing infection than bacteria at presentation. Improved culture-independent techniques, such as protein biomarker identification and 16S rRNA gene sequencing, will improve diagnostic and prognostic abilities and give greater power to future studies that investigate these issues.

As biofilms continue to complicate open fracture care by establishing persistent infections of implanted hardware and host tissue substrates, researchers will likely continue to develop new methods to prevent and eradicate them. Promising methods include the coating of hardware devices with actively released antibiotics, antisense molecules, quorum sensing inhibitors, bacteriophages, and immune-system-stimulating cytokines. Effective alternative methods to antibiotic treatment for established biofilm infections, such as bacteriophage therapy, need to be further explored because antibiotics poorly penetrate biofilms.

4.6 Summary

Open fracture wounds are at an increased risk for developing infections and other related complications. Current treatment paradigms aim to minimize infectious risks by effectively categorizing the wounds, treating the patients with prophylactic antibiotics, effectively debriding and irrigating the wounds, and appropriately fixing the fractures. While these treatment methods are well established in modern practice, many therapeutic details remain a point of discussion, such as the prophylactic benefits of Gram negative antibiotics.

Despite the efficacy of contemporary treatment paradigms, current research is continuing to address the deficiencies in current care methods. This research includes the use of culture-independent techniques, including bacterial DNA sequencing and protein biomarker detection, for assessing open fracture contamination or infection. Improved methods are also being developed for the removal of contamination and treatment of infection. Biofilm formation on fracture fixation hardware is a major concern, and techniques are being developed to prevent these infections, including various hardware coating techniques. One such coating technique aims to stimulate the antibacterial immune response, and this is also being developed as a compound to be administered with antibiotics to improve their overall efficacy. Additionally, treatments involving immune system stimulation are being developed to address the local deficient immune responses of open fractures.

4.7 Acknowledgments and Funding Sources

The authors thank the members of the Elizabeth Grice Laboratory, as well as the members of the Department of Orthopaedic Surgery at the Hospital of University of Pennsylvania, for their underlying contributions. G.D.H. is supported by the Department of Defense through the National Defense Science and Engineering Graduate Fellowship Program.

4.8 Author Disclosure and Ghostwriting

The authors do not declare any conflicts of interest. The authors do not declare any ghostwriter contributions.

4.9 Figures

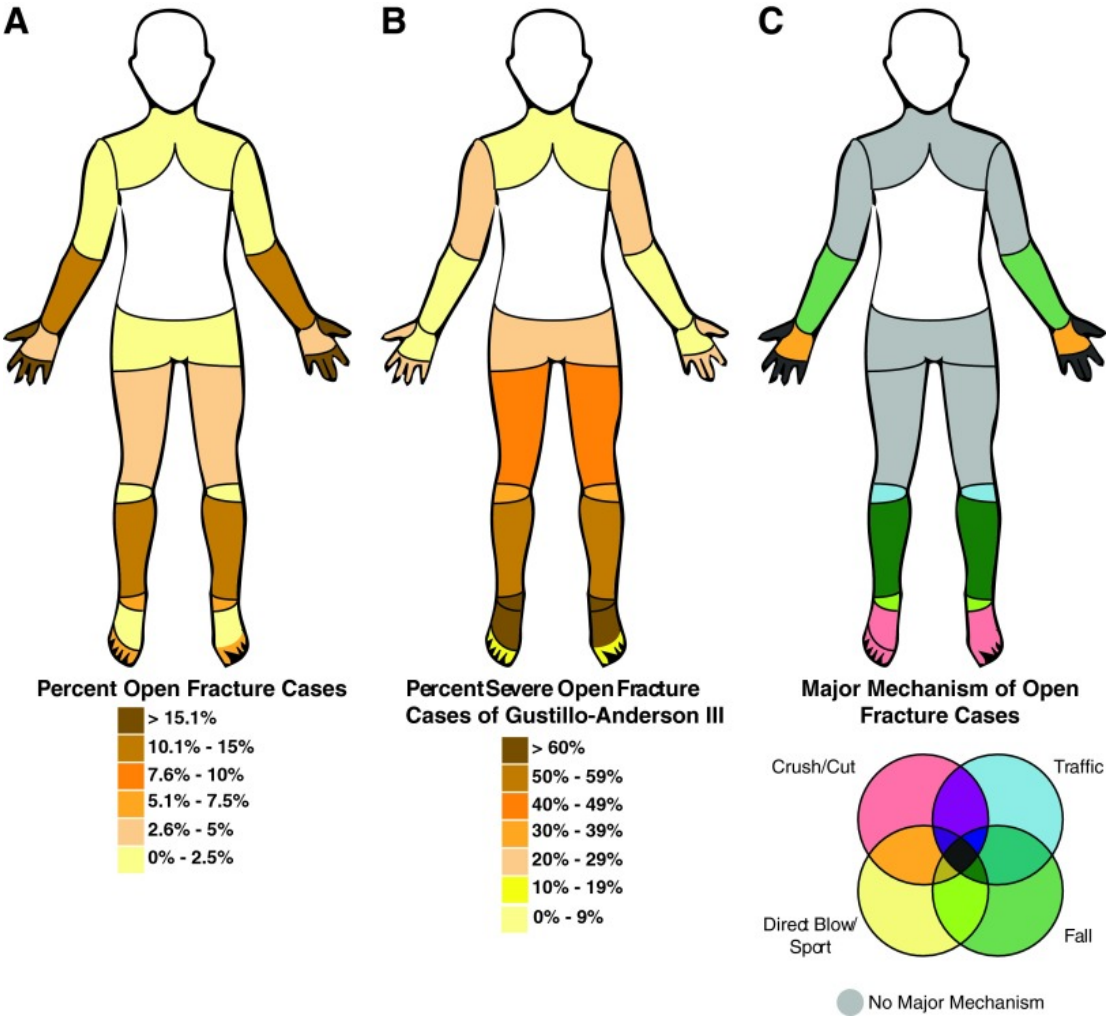

Figure 1 Open fracture rate, severity, and mechanistic cause statistics.

Open fracture rates and statistics, grouped by anatomical site, from a recent report by Court-Brown et al.1 The information represents a collection of 2386 open fracture cases recorded at the Royal Infirmary of Edinburgh between 1995 and 2009. The data suggest the majority of open fracture cases occurred on the distant extremities (A). The most severe open fractures (GA Type III) occurred on the lower extremities, especially the lower legs and feet (B). The distant extremities were characterized by major open fracture mechanisms, which have been grouped into four categories for easier visualization (C).

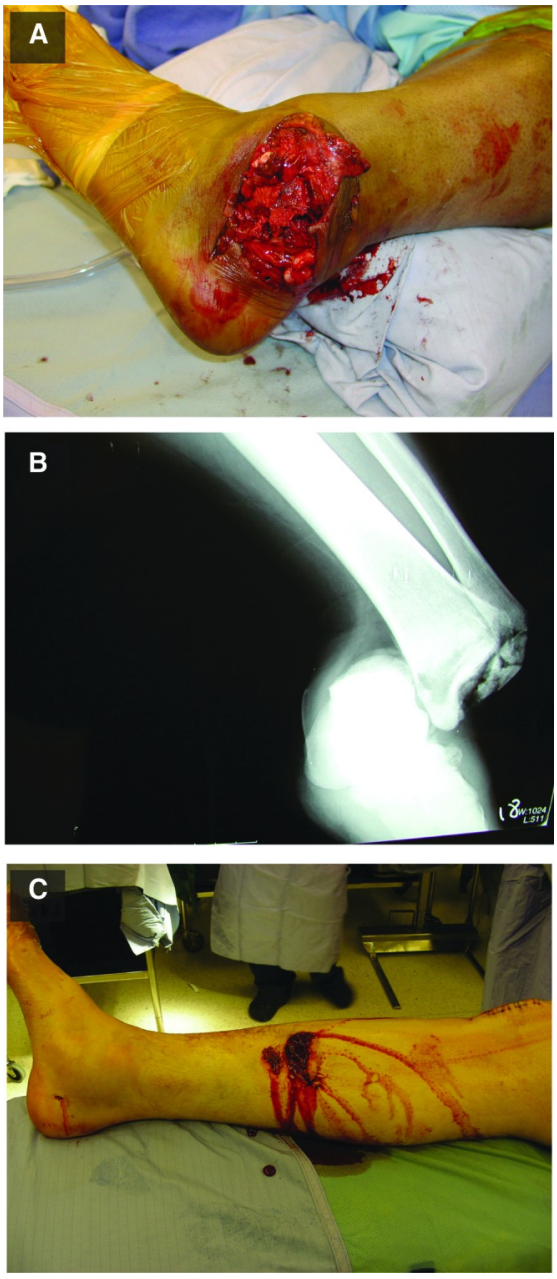

Figure 2 Examples of Gustilo–Anderson wound severities.

An example of a Gustilo–Anderson Type III open fracture that exhibits extensive soft tissue damage with minimal coverage (A). (B) An X-ray image of the wound in (A). (C) A Type II open fracture with minimal soft tissue damage.

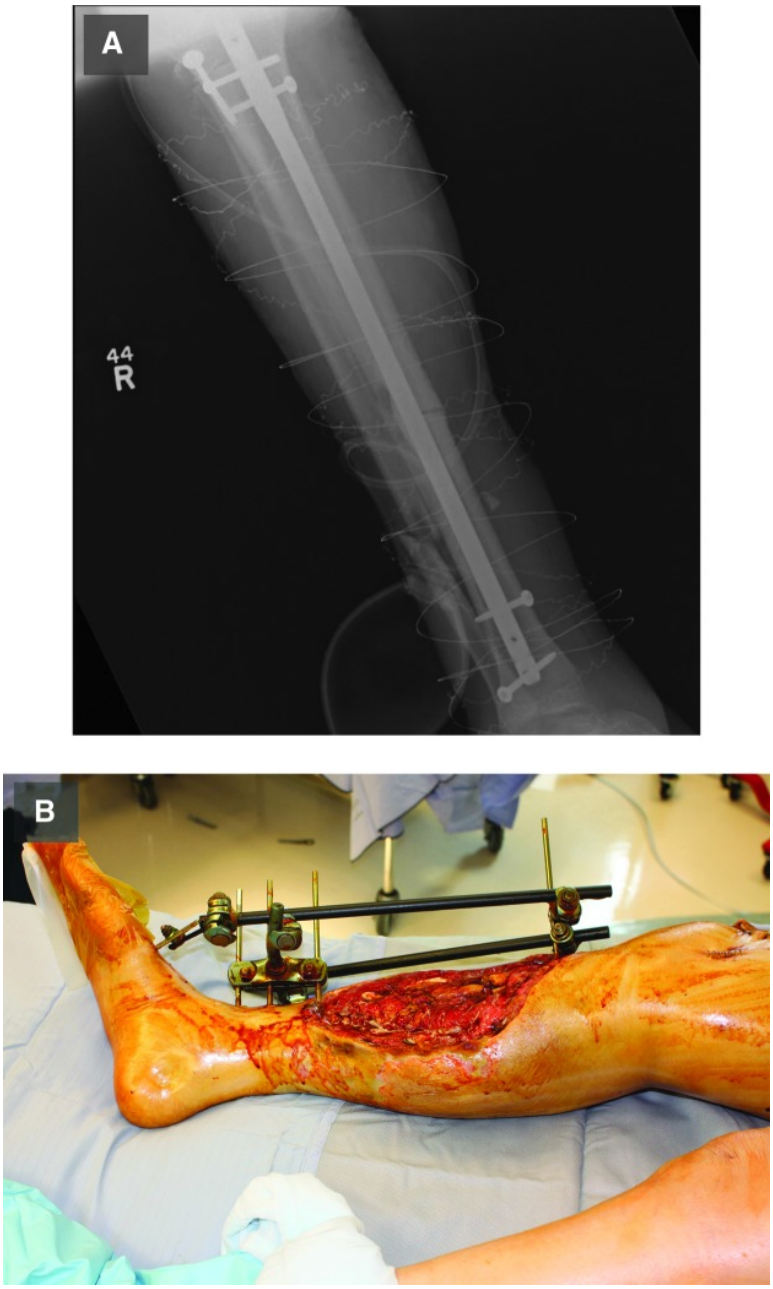

Figure 3 Example of fracture fixation techniques and hardware.

An example of an intramedullary nail used to fix an open diaphyseal tibia fracture (A). The fibula was also fractured but it was not fixed because it is not a weight-bearing bone (A). An example of an external fixation device being used to fix an open tibial fracture (B).

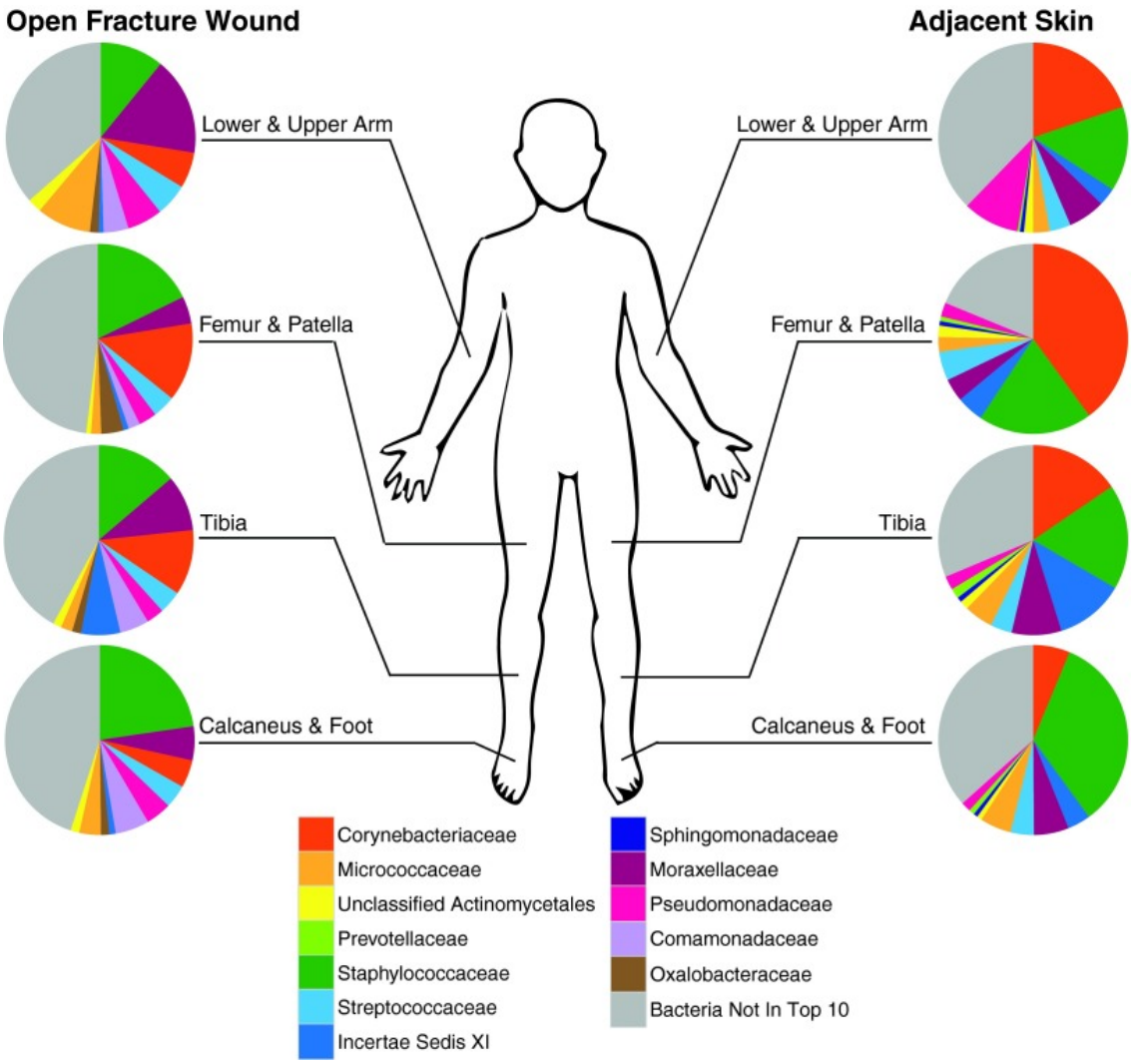

Figure 4 Bacterial communities associated with open fractures at emergency room presentation.

The bacterial communities of open fracture wounds (left) and their corresponding adjacent, unaffected skin (right), as reported by Hannigan et al [52]. The communities were grouped into four anatomical categories. The top 10 bacterial families, calculated as median relative abundance across all samples, were calculated for the wound and skin groups. The bacterial communities upon patient presentation to the emergency room are shown. The skin communities are primarily dominated by Corynebacteriaceae and Staphylococcaceae, while the wound communities are less dominated by these or other bacteria. The wound and skin communities differ from each other at

the same anatomical locations, and the different anatomically located communities also differ within the wound and skin categories. The bacteria labels are listed in the legend near the figure bottom.

4.10 Tables

Classification Score	*Wound Size (cm)*	*Soft Tissue Damage*
Type I	<1	Minimal
Type II	>1	Minimal
Type III		
A	>1	Extensive damage with adequate coverage
B	>1	Extensive damage with inadequate coverage
C	>1	Extensive damage, inadequate coverage, and extensive arterial and vascular damage

Table 1

The Gustilo and Anderson classification scheme. Adapted from Melvin *et al.*9 and Rüedi and Murphy.10

Grade	*Soft Tissue Damage*	*Mechanism of Injury*	*Associated Contamination*	*Other Considerations*
Grade 1	Minimal	Indirect trauma	Minimal	Small puncture wound without skin contusion
Grade 2	Moderate	Direct trauma	Moderate	Small skin and soft tissue contusions
Grade 3	Extensive	Farming accidents, gunshot wounds, and compartment syndrome	Extensive	Arterial and/or neural injuries
Grade 4	Extensive	Subtotal and total amputation	Extensive	

Table 2

Tscherne classification scheme for open fractures. Adapted from Rüedi and Murphy10 and Moore D (www.orthobullets.com/trauma/1002/tscherne-classification).

Affected Tissue	*Classification Score*	*Damage Severity*
Skin	IO 1	Minimal with skin breakage from inside out
	IO 2	Minimal with skin breakage from outside in (<5 cm)
	IO 3	Moderate with skin breakage from outside in (>5 cm)
	IO 4	Extensive with full-thickness contusion, abrasion, open degloving, and skin loss
	IO 5	Extensive with severe degloving
Muscle/Tendon	MT 1	None
	MT 2	Minimal and local
	MT 3	Moderate
	MT 4	Extensive with muscle defect and tendon laceration
	MT 5	Extensive with wide zone of injury and compartment syndrome
Neurovasculature	NV 1	None
	NV 2	Isolated
	NV 3	Local
	NV 4	Extensive with vascular injury
	NV 5	Extensive with subtotal or total amputation

Table 3

AO classification of open fractures. Information from Melvin et al.9 and Rüedi and Murphy.10

AO, Arbeitsgemeinschaft für Osteosynthesefragen; IO, open integument.

	Method	*References*	*Details*
Contamination	Surveillance culturing	50	Isolation of microbes on artificial media, followed by identification using biochemical and molecular techniques. Surveillance cultures are thought to have little value in predicting infections.
	DNA sequencing	23	The sequencing of taxonomically/phylogenetically informative genes to allow for bacterial community identification. This method is in development and is not used clinically.
	qPCR detection	4,23	Quantification of microbial load using qPCR of conserved microbial genes. This method is not widely used for clinical prophylaxis of open fractures.
	Host protein biomarker identification	59	Some host proteins have been identified as potential biomarkers for predicting open fracture healing complications. Many of these proteins have established roles in the immune response, and may be clinically useful upon further investigation.
Infection	Immune-related marker	65,66	Identification of host biomarkers, such as secondary rises in C-reactive protein levels, provides good support that a wound or implanted device is infected.
	Quantitative and qualitative culturing		Culturing of tissue samples can provide insight into what potentially pathogenic microbes are present at a site suspected of being infected, and can be used to predict the most prominent microbes. This can help in the prediction of the infecting organism itself.
	Histopathology		Histopathological examination of tissue sample infiltration by inflammatory cells can provide evidence for infection.
	X-Ray imaging		X-ray evidence of hardware implant nail loosening or widening of the fracture gap both suggest the presence of infection, although this alone is not definitive.
	Nuclear imaging		Nuclear imaging showing a lack of ^{99m}Tc accumulation, which indicates dead or devascularized bone, is suggestive of infection.
	Computed tomography		MRI allows for detection of soft tissue abnormalities, but is not effective for areas around metallic implants. PET and PET-CT scanning systems are able to address this deficiency in MRI approaches by allowing for assessment of accumulation of compounds, such as FDG, which indicate potential infection.

Table 4

Methods for detecting open fracture microbial contamination/infection. FDG, fluorine 18 fluorodeoxyglucose; MRI, magnetic resonance imaging; PET, positron emission tomography; PET-CT, positron emission tomography-computed tomography; qPCR, quantitative polymerase chain reaction.

4.11 References

1. Court-Brown CM, Bugler KE, Clement ND, Duckworth AD, McQueen MM: **The epidemiology of open fractures in adults. A 15-year review.** *Injury* 2012, **43:**891-897.
2. Patzakis MJ, Harvey JP, Jr., Ivler D: **The role of antibiotics in the management of open fractures.** *J Bone Joint Surg Am* 1974, **56:**532-541.
3. MacKenzie EJ, Jones AS, Bosse MJ, Castillo RC, Pollak AN, Webb LX, Swiontkowski MF, Kellam JF, Smith DG, Sanders RW, et al: **Health-care costs associated with amputation or reconstruction of a limb-threatening injury.** *J Bone Joint Surg Am* 2007, **89:**1685-1692.
4. Gille J, Wallstabe S, Schulz A-P, Paech A, Gerlach U: **Is non-union of tibial shaft fractures due to nonculturable bacterial pathogens? A clinical investigation using PCR and culture techniques.** *Journal of orthopaedic surgery and research* 2012, **7:**20.
5. Gustilo RB, Anderson JT: **Prevention of infection in the treatment of one thousand and twenty-five open fractures of long bones: retrospective and prospective analyses.** *J Bone Joint Surg Am* 1976, **58:**453-458.
6. Gustilo RB, Mendoza RM, Williams DN: **Problems in the management of type III (severe) open fractures: a new classification of type III open fractures.** *J Trauma* 1984, **24:**742-746.
7. Patzakis MJ, Wilkins J: **Factors influencing infection rate in open fracture wounds.** *Clin Orthop Relat Res* 1989**:**36-40.
8. Oestern HJ, Tscherne H: **Pathophysiology and classification of soft tissue injuries associated with fractures.** *Fractures with Soft Tissue Injuries* 1984**:**1-9.
9. Melvin JS, Dombroski DG, Torbert JT, Kovach SJ, Esterhai JL, Mehta S: **Open tibial shaft fractures: I. Evaluation and initial wound management.** *J Am Acad Orthop Surg* 2010, **18:**10-19.
10. Rüedi TP, Murphy WM: *AO principles of fracture management.* Stuttgart ; New York

Davos Platz, Switzerland: Thieme ;

AO Pub.; 2000.
11. Braun R, Enzler MA, Rittmann WW: **A double-blind clinical trial of prophylactic cloxacillin in open fractures.** *J Orthop Trauma* 1987, **1:**12-17.
12. Bergman BR: **Antibiotic prophylaxis in open and closed fractures: a controlled clinical trial.** *Acta Orthop Scand* 1982, **53:**57-62.
13. Hauser CJ, Adams CA, Jr., Eachempati SR, Council of the Surgical Infection S: **Surgical Infection Society guideline: prophylactic antibiotic use in open fractures: an evidence-based guideline.** *Surg Infect (Larchmt)* 2006, **7:**379-405.
14. Saveli CC, Belknap RW, Morgan SJ, Price CS: **The role of prophylactic antibiotics in open fractures in an era of community-acquired methicillin-resistant Staphylococcus aureus.** *Orthopedics* 2011, **34:**611-616; quiz 617.
15. Rohde RE, Denham R, Brannon A: **Methicillin resistant Staphylococcus aureus: carriage rates and characterization of students in a Texas university.** *Clin Lab Sci* 2009, **22:**176-184.
16. Shukla S, Nixon M, Acharya M, Korim MT, Pandey R: **Incidence of MRSA surgical-site infection in MRSA carriers in an orthopaedic trauma unit.** *J Bone Joint Surg Br* 2009, **91:**225-228.

17. Schweizer M, Perencevich E, McDanel J, Carson J, Formanek M, Hafner J, Braun B, Herwaldt L: **Effectiveness of a bundled intervention of decolonization and prophylaxis to decrease Gram positive surgical site infections after cardiac or orthopedic surgery: systematic review and meta-analysis.** *BMJ* 2013, **346:**f2743.
18. Hannigan GD, Pulos N, Grice EA, Mehta S: **Current Concepts and Ongoing Research in the Prevention and Treatment of Open Fracture Infections.** *Adv Wound Care (New Rochelle)* 2015, **4:**59-74.
19. Mody RM, Zapor M, Hartzell JD, Robben PM, Waterman P, Wood-Morris R, Trotta R, Andersen RC, Wortmann G: **Infectious complications of damage control orthopedics in war trauma.** *J Trauma* 2009, **67:**758-761.
20. Glass GE, Barrett SP, Sanderson F, Pearse MF, Nanchahal J: **The microbiological basis for a revised antibiotic regimen in high-energy tibial fractures: preventing deep infections by nosocomial organisms.** *J Plast Reconstr Aesthet Surg* 2011, **64:**375-380.
21. Johnson EN, Burns TC, Hayda RA, Hospenthal DR, Murray CK: **Infectious complications of open type III tibial fractures among combat casualties.** *Clin Infect Dis* 2007, **45:**409-415.
22. Snitkin ES, Zelazny AM, Montero CI, Stock F, Mijares L, Program NCS, Murray PR, Segre JA: **Genome-wide recombination drives diversification of epidemic strains of Acinetobacter baumannii.** *Proc Natl Acad Sci U S A* 2011, **108:**13758-13763.
23. Lockhart SR, Abramson MA, Beekmann SE, Gallagher G, Riedel S, Diekema DJ, Quinn JP, Doern GV: **Antimicrobial resistance among Gram-negative bacilli causing infections in intensive care unit patients in the United States between 1993 and 2004.** *J Clin Microbiol* 2007, **45:**3352-3359.
24. Gaynes R, Edwards JR, National Nosocomial Infections Surveillance S: **Overview of nosocomial infections caused by gram-negative bacilli.** *Clin Infect Dis* 2005, **41:**848-854.
25. Centers for Disease C, Prevention: **Acinetobacter baumannii infections among patients at military medical facilities treating injured U.S. service members, 2002-2004.** *MMWR Morb Mortal Wkly Rep* 2004, **53:**1063-1066.
26. Zalavras CG, Patzakis MJ, Holtom PD, Sherman R: **Management of open fractures.** *Infect Dis Clin North Am* 2005, **19:**915-929.
27. Dellinger EP, Caplan ES, Weaver LD, Wertz MJ, Droppert BM, Hoyt N, Brumback R, Burgess A, Poka A, Benirschke SK, et al.: **Duration of preventive antibiotic administration for open extremity fractures.** *Arch Surg* 1988, **123:**333-339.
28. Friedrich PL: **Die aseptische versorgung frischer wundern.** *Arch Klin Chir* 1898**:**288-310.
29. Schenker ML, Yannascoli S, Baldwin KD, Ahn J, Mehta S: **Does timing to operative debridement affect infectious complications in open long-bone fractures? A systematic review.** *J Bone Joint Surg Am* 2012, **94:**1057-1064.
30. Anglen JO: **Wound irrigation in musculoskeletal injury.** *J Am Acad Orthop Surg* 2001, **9:**219-226.
31. Penn-Barwell JG, Murray CK, Wenke JC: **Comparison of the antimicrobial effect of chlorhexidine and saline for irrigating a contaminated open fracture model.** *J Orthop Trauma* 2012, **26:**728-732.

32. Anglen JO: **Comparison of soap and antibiotic solutions for irrigation of lower-limb open fracture wounds. A prospective, randomized study.** *J Bone Joint Surg Am* 2005, **87:**1415-1422.
33. Anglen J, Apostoles PS, Christensen G, Gainor B, Lane J: **Removal of surface bacteria by irrigation.** *J Orthop Res* 1996, **14:**251-254.
34. Bhandari M, Adili A, Schemitsch EH: **The efficacy of low-pressure lavage with different irrigating solutions to remove adherent bacteria from bone.** *J Bone Joint Surg Am* 2001, **83-A:**412-419.
35. Dirschl DR, Duff GP, Dahners LE, Edin M, Rahn BA, Miclau T: **High pressure pulsatile lavage irrigation of intraarticular fractures: effects on fracture healing.** *J Orthop Trauma* 1998, **12:**460-463.
36. Hassinger SM, Harding G, Wongworawat MD: **High-pressure pulsatile lavage propagates bacteria into soft tissue.** *Clin Orthop Relat Res* 2005, **439:**27-31.
37. Boyd JI, 3rd, Wongworawat MD: **High-pressure pulsatile lavage causes soft tissue damage.** *Clin Orthop Relat Res* 2004**:**13-17.
38. Melvin JS, Dombroski DG, Torbert JT, Kovach SJ, Esterhai JL, Mehta S: **Open tibial shaft fractures: II. Definitive management and limb salvage.** *J Am Acad Orthop Surg* 2010, **18:**108-117.
39. Esposito J, Schemitsch EH, Saccone M, Sternheim A, Kuzyk PR: **External fixation versus open reduction with plate fixation for distal radius fractures: a meta-analysis of randomised controlled trials.** *Injury* 2013, **44:**409-416.
40. Canadian Orthopaedic Trauma S: **Nonunion following intramedullary nailing of the femur with and without reaming. Results of a multicenter randomized clinical trial.** *J Bone Joint Surg Am* 2003, **85-A:**2093-2096.
41. Wiss DA, Stetson WB: **Unstable fractures of the tibia treated with a reamed intramedullary interlocking nail.** *Clin Orthop Relat Res* 1995**:**56-63.
42. Keating JF, O'Brien PJ, Blachut PA, Meek RN, Broekhuyse HM: **Locking intramedullary nailing with and without reaming for open fractures of the tibial shaft. A prospective, randomized study.** *J Bone Joint Surg Am* 1997, **79:**334-341.
43. Finkemeier CG, Schmidt AH, Kyle RF, Templeman DC, Varecka TF: **A prospective, randomized study of intramedullary nails inserted with and without reaming for the treatment of open and closed fractures of the tibial shaft.** *J Orthop Trauma* 2000, **14:**187-193.
44. Hutchinson AJ, Frampton AE, Bhattacharya R: **Operative fixation for complex tibial fractures.** *Ann R Coll Surg Engl* 2012, **94:**34-38.
45. Valenziano CP, Chattar-Cora D, O'Neill A, Hubli EH, Cudjoe EA: **Efficacy of primary wound cultures in long bone open extremity fractures: are they of any value?** *Arch Orthop Trauma Surg* 2002, **122:**259-261.
46. Lee J: **Efficacy of cultures in the management of open fractures.** *Clin Orthop Relat Res* 1997**:**71-75.
47. Fischer MD, Gustilo RB, Varecka TF: **The timing of flap coverage, bone-grafting, and intramedullary nailing in patients who have a fracture of the tibial shaft with extensive soft-tissue injury.** *J Bone Joint Surg Am* 1991, **73:**1316-1322.
48. Burns TC, Stinner DJ, Mack AW, Potter BK, Beer R, Eckel TT, Possley DR, Beltran MJ, Hayda RA, Andersen RC, et al: **Microbiology and injury characteristics in severe open tibia fractures from combat.** *Journal of Trauma and Acute Care Surgery* 2012, **72:**1062-1067.

49. Sen RK, Murthy N, Gill SS, Nagi ON: **Bacterial load in tissues and its predictive value for infection in open fractures.** *J Orthop Surg (Hong Kong)* 2000, **8:**1-5.
50. Cooney WP, 3rd, Fitzgerald RH, Jr., Dobyns JH, Washington JA, 2nd: **Quantitative wound cultures in upper extremity trauma.** *J Trauma* 1982, **22:**112-117.
51. Merritt K: **Factors increasing the risk of infection in patients with open fractures.** *J Trauma* 1988, **28:**823-827.
52. Hannigan GD, Hodkinson BP, McGinnis K, Tyldsley AS, Anari JB, Horan AD, Grice EA, Mehta S: **Culture-independent pilot study of microbiota colonizing open fractures and association with severity, mechanism, location, and complication from presentation to early outpatient follow-up.** *J Orthop Res* 2014, **32:**597-605.
53. Hannigan GD, Grice EA: **Microbial ecology of the skin in the era of metagenomics and molecular microbiology.** *Cold Spring Harb Perspect Med* 2013, **3:**a015362.
54. Cho I, Blaser MJ: **The human microbiome: at the interface of health and disease.** *Nat Rev Genet* 2012, **13:**260-270.
55. Chromy BA, Eldridge A, Forsberg JA, Brown TS, Kirkup BC, Jaing C, Be NA, Elster E, Luciw PA: **Wound outcome in combat injuries is associated with a unique set of protein biomarkers.** *J Transl Med* 2013, **11:**281.
56. Grice EA, Snitkin ES, Yockey LJ, Bermudez DM, Program NCS, Liechty KW, Segre JA: **Longitudinal shift in diabetic wound microbiota correlates with prolonged skin defense response.** *Proc Natl Acad Sci U S A* 2010, **107:**14799-14804.
57. Ricklin D, Hajishengallis G, Yang K, Lambris JD: **Complement: a key system for immune surveillance and homeostasis.** *Nat Immunol* 2010, **11:**785-797.
58. Zalavras CG, Patzakis MJ, Holtom P: **Local antibiotic therapy in the treatment of open fractures and osteomyelitis.** *Clin Orthop Relat Res* 2004**:**86-93.
59. Costerton JW, Post JC, Ehrlich GD, Hu FZ, Kreft R, Nistico L, Kathju S, Stoodley P, Hall-Stoodley L, Maale G, et al: **New methods for the detection of orthopedic and other biofilm infections.** *FEMS Immunol Med Microbiol* 2011, **61:**133-140.
60. Trampuz A, Widmer AF: **Infections associated with orthopedic implants.** *Curr Opin Infect Dis* 2006, **19:**349-356.
61. Trampuz A, Zimmerli W: **Diagnosis and treatment of infections associated with fracture-fixation devices.** *Injury* 2006, **37 Suppl 2:**S59-66.
62. Vergidis P, Patel R: **Novel approaches to the diagnosis, prevention, and treatment of medical device-associated infections.** *Infect Dis Clin North Am* 2012, **26:**173-186.
63. Wolcott R, Costerton JW, Raoult D, Cutler SJ: **The polymicrobial nature of biofilm infection.** *Clin Microbiol Infect* 2013, **19:**107-112.
64. Williams DL, Haymond BS, Beck JP, Savage PB, Chaudhary V, Epperson RT, Kawaguchi B, Bloebaum RD: **In vivo efficacy of a siliconecationic steroid antimicrobial coating to prevent implant-related infection.** *Biomaterials* 2012, **33:**8641-8656.
65. Boyce BM, Lindsey BA, Clovis NB, Smith ES, Hobbs GR, Hubbard DF, Emery SE, Barnett JB, Li B: **Additive effects of exogenous IL-12 supplementation and antibiotic treatment in infection prophylaxis.** *J Orthop Res* 2012, **30:**196-202.
66. Li B, Jiang B, Boyce BM, Lindsey BA: **Multilayer polypeptide nanoscale coatings incorporating IL-12 for the prevention of biomedical device-associated infections.** *Biomaterials* 2009, **30:**2552-2558.

CHAPTER 5 – Culture-independent pilot study of microbiota colonizing open fractures and association with severity, mechanism, location, and complication from presentation to early outpatient follow-up

The contents of this chapter have been published as:

Hannigan GD, Hodkinson BP, McGinnis K, Tyldsley AS, Anari JB, Horan AD, Grice EA, Mehta S. 2014. Culture-independent pilot study of microbiota colonizing open fractures and association with severity, mechanism, location, and complication from presentation to early outpatient follow-up. J Orthop Res. 32(4):597-605. doi: 10.1002/jor.22578.

5.1 Abstract

Precise identification of bacteria associated with post-injury infection, co-morbidities, and outcomes could have a tremendous impact in the management and treatment of open fractures. We characterized microbiota colonizing open fractures using culture-independent, high-throughput DNA sequencing of bacterial 16S ribosomal RNA genes, and analyzed those communities with respect to injury mechanism, severity, anatomical site, and infectious complications. Thirty subjects presenting to the Hospital of the University of Pennsylvania for acute care of open fractures were enrolled in a prospective cohort study. Microbiota was collected from wound center and adjacent skin upon presentation to the emergency department, intraoperatively, and at two outpatient follow-up visits at approximately 25 and 50 days following initial presentation. Bacterial community composition and diversity colonizing open fracture wounds became increasingly similar to adjacent skin microbiota with healing. Mechanism of injury, severity, complication, and location were all associated with various aspects of microbiota diversity and composition. The results of this pilot study demonstrate the diversity and dynamism of the open fracture microbiota, and their relationship to clinical variables. Validation of these

preliminary findings in larger cohorts may lead to the identification of microbiome-based biomarkers of complication risk and/or to aid in management and treatment of open fractures.

5.2 Introduction

Open fractures are characterized by soft tissue disruption at the fracture site increasing the risk of complications including infection, nonunion/malunion, and amputation. Infection risk increases with increasing injury severity and occurs up to 50% of the time when extensive soft tissue damage is involved, due to compromised vascularity among other factors [1]. Predicting which patients will have an infection remains difficult. Surveillance cultures at the time of presentation (before signs and symptoms) are generally thought to have little predictive value [1-3]. Reliable biomarkers to guide management and treatment of open fractures are needed. We hypothesized that microbiota colonizing open fractures during acute phases of injury, prior to clinical signs of infection, may be an information-rich read-out of the wound environment providing valuable insight into the mechanisms of impending complication.

Our bodies are colonized inside and out with myriad commensal microorganisms (the "microbiome") that have important roles in human health and disease. While many infectious states are seemingly caused by single microorganisms satisfying Koch's postulates, the role of the microbiome in modulating the host immune response and resistance to pathogenic and opportunistic microorganisms is increasingly evident. Microorganisms are exquisitely sensitive to their host environment, and likewise, the host immune response is calibrated to react rapidly and precisely to fluctuations in the microbiota. An intimate relationship between the microbiota and the underlying immune and defense response has been demonstrated in skin and cutaneous wounds [4-7]. In the setting of an open fracture, the skin microbiome is altered as a result of the

dramatic change in the local environment and contamination from the injury. Local microbial changes may have significant impact on both local and systemic host defenses, soft tissue healing, and, ultimately, clinical outcome.

Most reported studies characterizing bacteria colonizing and/or infecting open fractures rely on clinical culture-based methodology. Traditional hospital-based culture techniques, however, apply heavy selection pressure in favor of bacteria capable of thriving in restricted artificial growth conditions. The most commonly cultured bacteria in open fractures are Staphylococcus and Gram-negative isolates [8-10]. Advances in high-throughput DNA sequencing technology enable the study of the human microbiome via sequencing of the bacteria-specific 16S small subunit ribosomal RNA (rRNA) gene. These genomic approaches are increasingly accessible and provide greater resolution and precision by eliminating biases associated with culturing bacteria.

In this pilot study, the microbiome colonizing the open fracture and adjacent skin during the course of healing was evaluated. Sequencing of bacterial 16S rRNA genes was employed to define the composition and diversity of the microbiota in open fractures as healing progressed. Further analysis was done to assess potential correlations between the open fracture microbiome and clinical factors (location, mechanism, severity) and clinical outcomes.

5.3 Results

5.3.1 Composition of Microbiota Colonizing the Open Fracture Site and Adjacent Skin

The six bacterial genera present in >1% median relative abundance in the open fracture and adjacent skin were Staphylococcus, Corynebacterium, Streptococcus, Acinetobacter,

Anaerococcus, and Pseudomonas (Table 2). We also specifically examined the relative abundance of Propionibacterium, Escherichia, and unclassified Enterobacteriaceae (family containing the genera Klebsiella and Enterobacter), due to their known pathogenic potential in traumatic injuries [2, 11-15]. The relative abundance of Staphylococcus significantly increased and that of Pseudomonas significantly decreased in the wound center versus the adjacent skin during the time course ($p = 0.043$ and 0.039, respectively). Escherichia relative abundance significantly increased on the adjacent skin, but was unchanged in the wound ($p = 0.012$). At the ER time point, the genera Corynebacterium and Anaerococcus were significantly more abundant in the adjacent skin as compared to the wound, where Pseudomonas was significantly more abundant in the wound ($p = 0.004$, 0.008, and 0.036, respectively). Corynebacterium continued to be significantly higher in relative abundance on the skin compared to the wound even after DIC ($p = 0.030$).

5.3.2 Comparison of Findings From Culture-Independent and Culture-Dependent Methodologies

Wound cultures were obtained for 14 of the 30 subjects at the time of presentation to the ER. 2/14 (13%) were culture positive for bacteria, with one being culture positive for Stenotrophomonas maltophilia and one was culture positive for Enterobacter cloacae. 16S rRNA profiling indicated the presence of Stenotrophomonas in the wound from which Stenotrophomonas maltophilia was cultured. We did not detect Enterobacter in the open fracture that cultured positive for Enterobacter cloacae, likely due to limitations of 16S rRNA sequence-based identification and taxonomic classification. However, we did detect unclassified Enterobacteriaceae, which is the family-level taxon that encompasses Enterobacter species.

Of the seven subjects in this study that presented with eventual complication, cultures were obtained as standard of care for three of the subjects at the time of complication. Two of the three subjects were culture positive for Staphylococcus (one coagulase-negative and one MRSA) at the time of surgery for nonunion and multiple debridement surgeries, respectively. We detected Staphylococcus by 16S rRNA sequencing in all skin and wound samples at all time points of sampling for these subjects. The third subject developed an infection that was culture positive for Staphylococcus aureus, Peptostreptococcus, and Enterococcus. At the ER time point, we detected Peptostreptococcus in skin and wound samples and Enterococcus on the skin. Enterococcus was detected on the skin at all time points and Peptostreptococcus was detected in skin and wound samples at 2nd OP. These findings suggest that the eventual type of bacteria implicated in complication by cultures may be present as early as presentation to the ER, and may result from contamination from skin microbiota or be present in the wound itself.

5.3.3 Dynamic Microbial Diversity of Open Fracture and Convergence With Adjacent Skin Microbiota

To gain an overall view of bacterial community structures changing over time, the beta diversity of the open fracture wound to the corresponding adjacent skin at each time point was compared. Beta diversity was calculated for each pair of samples using the Bray–Curtis metric, which takes into account the number of shared species-level OTUs and their abundance. PCoA plots were used to visualize the shared diversity of wound and the adjacent skin at presentation to the ER (Fig. 1A), at 1st OP (Fig. 1B), and at 2nd OP (Fig. 1C). Progressively, skin and wound communities converged, becoming increasingly similar to each other at each subsequent time point, as measured by Median Intersample Dissimilarity (MID), where a higher MID value indicates greater dissimilarity. ER, 1st OP, and 2nd OP MID values were 0.690, 0.674, and 0.445, respectively. Significant differences between skin and wound microbiomes only existed at the ER

time point ($p = 0.039$; $R = 0.124$; Fig. 1). At the latter two time points, wound and skin bacterial community structures are indistinguishable by the metrics employed. Given that 6/21 and 5/15 samples analyzed at 1st and 2nd OP, respectively were considered healed at those time points, convergence of wound microbiota with the skin microbiota would be expected.

Alpha diversity of open fracture microbiota was measured by the number of observed species-level OTUs and Faith's Phylogenetic Diversity index (Faith's PD), a metric that takes into account phylogenetic branch length in addition to the number of OTUs present in a sample. These analyses revealed significantly decreased alpha diversity in the wound compared to the skin at presentation to the ER ($p = 0.019$ and $p = 0.006$ for observed OTUs and Faith's PD, respectively; Fig. 2A and B). There was also a significant decrease in adjacent skin alpha diversity at the first clinical follow-up compared to when the patient presented to the ER ($p = 0.011$ and $p = 0.003$ for observed OTUs and Faith's PD, respectively). We independently examined total bacterial load by quantitative PCR of the 16S rRNA gene. We did not observe significant differences between ER, OR, 1st OP, and 2nd OP time points, or between the wound center & adjacent skin (Fig. 2C).

Because of the synergistic role that Gram-positive and -negative organisms have in forming biofilms in wounds and on orthopaedic devices [13], we compared the relative abundances of Gram-positive and -negative bacteria (Fig. 3). In the wound, relative abundances of each type of bacteria were approximately similar at presentation to the ER (Fig. 3A; $p = 0.908$), but Gram-positive bacteria were significantly more abundant on the skin than Gram-negative bacteria at the same time point (Fig. 3B; $p = 1.73 \times 10{-}11$). These differences were not detectable following DIC. However, at the 1st and 2nd OP time points, both the skin ($p = 0.003$ and $p = 2.58 \times 10{-}8$,

respectively) and wound ($p = 0.016$ and $p = 3.51 \times 10{-}6$, respectively) harbored greater relative abundance of Gram-positive bacteria, indicating a return to the original skin-like state.

5.3.4 Injury Mechanism, Location, Severity, and Complication Are Associated With Open Fracture Microbiota

We next analyzed open fracture and adjacent skin microbiomes with respect to clinical factors. We selected four variables noted at time of enrollment or in follow up: mechanism, location, progression to infectious complication, and Gustilo–Anderson classification. When examining mechanism and wound severity with respect to colonizing microbiota, alpha diversity, as measured by Faith's PD (Fig. 4A and B) and observed species-level OTUs (data not shown), was not significantly different, nor was beta diversity as measured by the Bray–Curtis metric (data not shown). However, when analyzing the top six genera present in >1% total abundance and those genera of interest (Table 3), we found that Corynebacterium relative abundance was significantly greater and unclassified Enterobactericeae relative abundance was significantly lesser in penetrating wounds compared to blunt wounds at the 1st OP time point ($p = 0.006$ and $p = 0.038$, respectively). At the 2nd OP time point, Pseudomonas relative abundance was significantly greater in penetrating wounds compared to blunt wounds ($p = 0.048$). Regarding severity, Type 1 fractures had increased relative abundance of Acinetobacter and decreased relative abundance of Propionibacterium compared to Type 3 injuries ($p = 0.015$ and $p = 0.038$, respectively; Table 3).

When analyzing microbiota with respect to development of complications, beta diversity, as measured by the Bray–Curtis metric, revealed significant differences in bacterial community structure ($p = 0.019$, $R = 0.176$) when comparing complicated to uncomplicated outcomes. We did not identify any significant changes in alpha diversity (Fig. 4C) nor in the specific genera we

selected for analysis, indicating that either rare bacteria present in <1% relative abundance or other undefined aspects of the microbiota are responsible for the change in community structure we observe when comparing the two groups.

Because skin microbial communities are known to differ by body site [16], we also selected wound location as a variable to analyze with respect to microbiomes. We grouped together open fractures of the upper extremities (humerus and ulna) and the lower extremities (femur, hip, tibia, fibula, foot). Bacterial community structure significantly differed when comparing beta diversity of the two groups using the Bray–Curtis metric ($p = 0.005$, $R = 0.300$). Lower extremity open fractures harbored greater alpha diversity than upper extremity fractures as measured by Faith's PD (Fig. 4D; $p = 0.036$) and observed species-level OTUs (data not shown; $p = 0.019$). When analyzing all time points, the genera Anaerococcus was significantly enriched in relative abundance in lower extremity compared to upper extremity open fractures (Table 3; $p = 0.015$).

5.4 Discussion

The findings from this pilot study using culture-independent, high-throughput sequencing based techniques, suggest that a great diversity of microbiota is present in open fractures. Follow-up studies, in larger cohorts and with more frequent sampling until healing is complete may provide more insight into the dynamic changes in the wound and skin microbiota, the association between the microbiota to clinical outcomes, and the potential predictive nature of colonizing bacteria. Similarly, based on a broader understanding of the microbiota, studies examining the role of early debridement, type and timing of antibiotic administration, and irrigation methods can be better designed. Concurrent molecular profiling of host genomic and expression profiles could further clarify mechanisms of infectious complications and the response to treatment.

Molecular techniques are a powerful tool in detecting bacteria. For example, biofilms, such as those that commonly grow on orthopaedic devices, are recalcitrant to culturing [17], suggesting the utility of DNA-based detection methods where biofilm is suspected. Commonly isolated organisms from orthopaedic devices are Staphylococcus, Pseudomonas, and Klebsiella [18, 19]. It is thought that polymicrobial biofilms, those consisting of both Gram-positive and Gram-negative bacteria, are more severe and recalcitrant to treatment [13]. Our findings reveal that, upon presentation to the ER, traumatic open fractures harbor a nearly equally abundant combination of commensal Gram-positive and -negative bacteria, though the skin is dominated by Gram-positive bacteria. The implications of this finding for biofilm formation are unclear, but it suggests that the substrates to nurture a polymicrobial biofilm are in place at the time of presentation. Early application of internal fixation may be at risk given the diversity of microbiome of an open fracture.

A novel aspect of this study was that we examined microbiomes of both the open fracture and the adjacent skin. The adjacent skin may be a source of contamination for open fractures. It may also provide a baseline for assessing microbiota of the open fracture. Together with our analysis of shared diversity at each time point, our data suggests that traumatic wound bacterial communities are least similar to healthy skin upon presentation to the ER, and as expected become more similar as healing progresses. Furthermore, mechanism of injury, location, and severity are associated with various aspects defining the colonizing microbiota, suggesting the need for different management techniques depending on the injury pattern, for example the difference between penetrating injuries and blunt force open fractures. The finding that open fractures that proceed to develop complications are associated with different microbial communities than those

that are complication-free indicates the potential prognostic value of 16S rRNA profiling for identifying those open fractures at risk for complication.

The limitations of our study are that we are in an urban setting with patients coming from the mid-Atlantic region. The local environmental microbiota may be different when comparing to other parts of the world, areas near open water, or wounds that occur on the battlefield across the world. Furthermore, we did not have a control group, which may have included a second individual not injured but in the vicinity of the injured patient. Hospital length of stay may also impact colonizing microbiota and progression to complication, and future studies in larger cohorts will need to take this potential nosocomial confounder into account. Lastly, some aspects of the analysis focused on those bacteria present in >1% relative abundance across the dataset. By including those species that have a known pathogenic potential and are clinically concerning, we attempted to address this.

Ultimately, this study reveals the complexity of the open fracture wound. The ramifications of improved understanding of the bacterial diversity, load, and noted taxa may have significant relevance to initial treatment, methods of monitoring, and clinical outcomes. Predictive modeling and biomarker panels may be the next step in further developing tools that can be applied clinically to decrease infection after open fractures.

5.5 Materials & Methods

5.5.1 Human Subjects Protections

Prior to study initiation, this protocol was reviewed and approved by the University of Pennsylvania School of Medicine Institutional Review Board. A modification of the informed consent process was approved for this investigation to enable sample collection under emergent conditions. Informed consent was obtained from all subjects enrolled in this study.

5.5.2 Sample Collection

Thirty open fracture patients from the Hospital of the University of Pennsylvania Orthopaedic Trauma and Fracture Service were recruited into the study. Characteristics of the patient population are summarized in Table 1. Using a Catch-All Sample Collection Swab (Epicentre), a microbiota sample was collected from the wound center and adjacent skin (5 cm away from the wound) of each subject at emergency room presentation (ER) prior to debridement, irrigation, and cleansing (DIC), and intraoperatively (OR) after DIC. Additional samples were collected at the first outpatient follow up visit (1st OP) and the outpatient visit closest to 28 days following 1st OP (2nd OP). At 1st OP and 2nd OP, 6/21 and 5/15 samples collected were from open fractures with healed soft tissue, respectively. Sample attrition, from the cohort of 30, occurred due to logistical issues in sample collection and attrition during trauma patient follow-up. Also, some samples did not amplify bacterial DNA in sufficient quantities to include in the analysis (see Supplementary Methods).

Negative control specimens were also collected by exposing swabs to room air and processing them alongside wound samples. Clinical, demographic, and behavioral information was collected for each participant. At initial presentation, each wound was classified according to the Gustilo–Anderson classification system [20], anatomic site, and injury mechanism. Complications were assessed as bivariates with any unplanned intervention in the post-operative period considered

positive (i.e., readmission, need for antibiotics, repeat debridement, or irrigation, soft tissue procedure).

5.5.3 DNA Isolation, Amplification, and Sequencing of 16S rRNA Genes

Detailed DNA extraction methodology is provided in the Supplemental Methods and has been previously described [21]. Detailed information on amplification procedure is also provided in Supplemental Methods. Sequencing of the V4 region was performed with the Illumina MiSeq system using 150 bp paired-end chemistry at the University of Pennsylvania Next Generation Sequencing Core. A total of 7,708,124 paired-end sequencing reads were included in the analysis, with a mean of 43,796 and a median of 30,048 sequences per sample.

5.5.4 Quantitative PCR (qPCR) of the 16S rRNA Gene

DNA from the swab extraction described above was used for qPCR-based bacterial load estimation. A portion of the 16S rRNA bacterial gene was amplified using the primers 533F (GTGCCAGCAGCCGCGGTAA) and 902R (GTCAATTCITTTGAGTTTYARYC) [22] on a ViiA7 platform (Applied Biosystems, Grand Island, NY). Each 10 µl reaction included 1 µl DNA, 5 µl 2× SYBR Green Master Mix (Invitrogen, Carlsbad, CA), and 0.1 µl of each 20 µM primer solution. Cycling conditions were 50°C (2 min), 95°C (10 min), and followed by 40 cycles of 95°C (15 sec) and 60°C (1 min). A standard curve was generated by amplifying serial dilutions of known concentrations of E. coli genomic DNA. Estimated 16S rRNA copy number and bacterial load were calculated as described previously [23].

5.5.5 16S rRNA Sequence Processing and Analyses

Details of 16S rRNA dataset processing and analyses are in the Supplemental Methods.

5.5.6 Statistical Analyses

The R statistical computing package was used for statistical analyses. Principle coordinates analysis (PCoA) plots were produced for visualizing distances between bacterial communities. ANOSIM tests were run to examine the relationship between sample groupings and overall community composition. p-values were calculated using 999 permutations. Wilcoxon rank-sum tests and Benjamini Hochberg false discovery rate (FDR) correction was applied to p-values to assess the significance of differences in: bacterial load, alpha diversity, and to test for significant associations in fracture characteristics with alpha diversity. Wilcoxon rank-sum tests were also used to assess the differences between taxon relative abundances of wound center and adjacent skin samples at specific time points and Kruskal–Wallis tests were used to examine taxon relative abundance changes across all time points for the following genera: (a) those with median relative abundances >1% in the entire dataset and (b) those that do not meet the 1% threshold but are designated as clinically relevant taxa of interest by the Department of Defense (i.e., Propionibacterium, Escherichia, Enterobacter, and Klebsiella). Because of inherent limitations of 16S rRNA-based taxonomic identification and classification, we could not resolve the genera Klebsiella and Enterobacter based on 16S rRNA sequence. We therefore include in these analyses the unclassified Enterobacteriaceae, which is the family-level taxon that includes the genera Klebsiella and Enterobacter.

5.6 Acknowledgments

This study was funded by the Orthopedic Trauma Association Grant #29 (S.M.), NIAMS/NIH R00 AR060873 (E.A.G.), internal funds from the University of Pennsylvania Perelman School of Medicine and the Department of Dermatology (E.A.G.), and the Department of Defense National

Defense Science and Engineering Graduate Fellowship (G.H.). We thank the Penn Orthopaedic Trauma and Fracture Service faculty, residents, and staff for their assistance with sample collection, and the study participants, without whom this study would not have been possible.

5.7 Figures

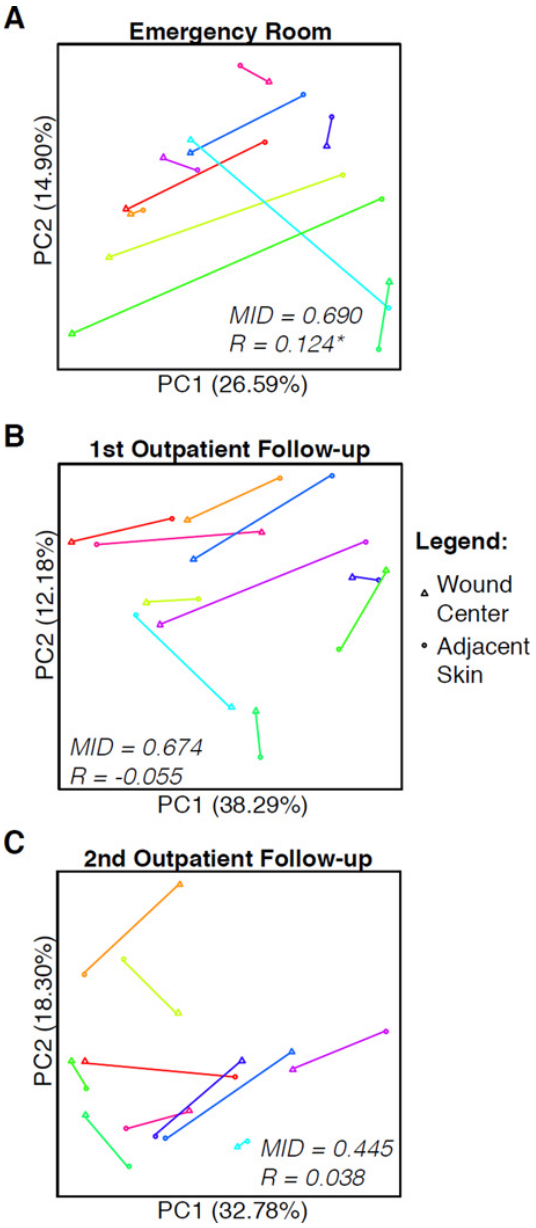

Figure 1 The changing relationships between open fracture wound and adjacent skin microbiota of 10 patients over time

PCoA plots representing the Bray-Curtis metric comparing beta diversity of open fracture and skin microbiota. Each color represents a different patient, while triangles and circles represent wound center and adjacent skin microbiota, respectively. Shown are the first two principle

coordinates and the percent variation explained by each principle coordinate is indicated in parentheses by the axis. The two samples (open fracture wound and adjacent skin) for each patient at a given time point are connected by a line. An ANOSIM test was used to examine the association between swab location and the overall community composition; this association is significant (at $P<0.05$) only for the ER time point.

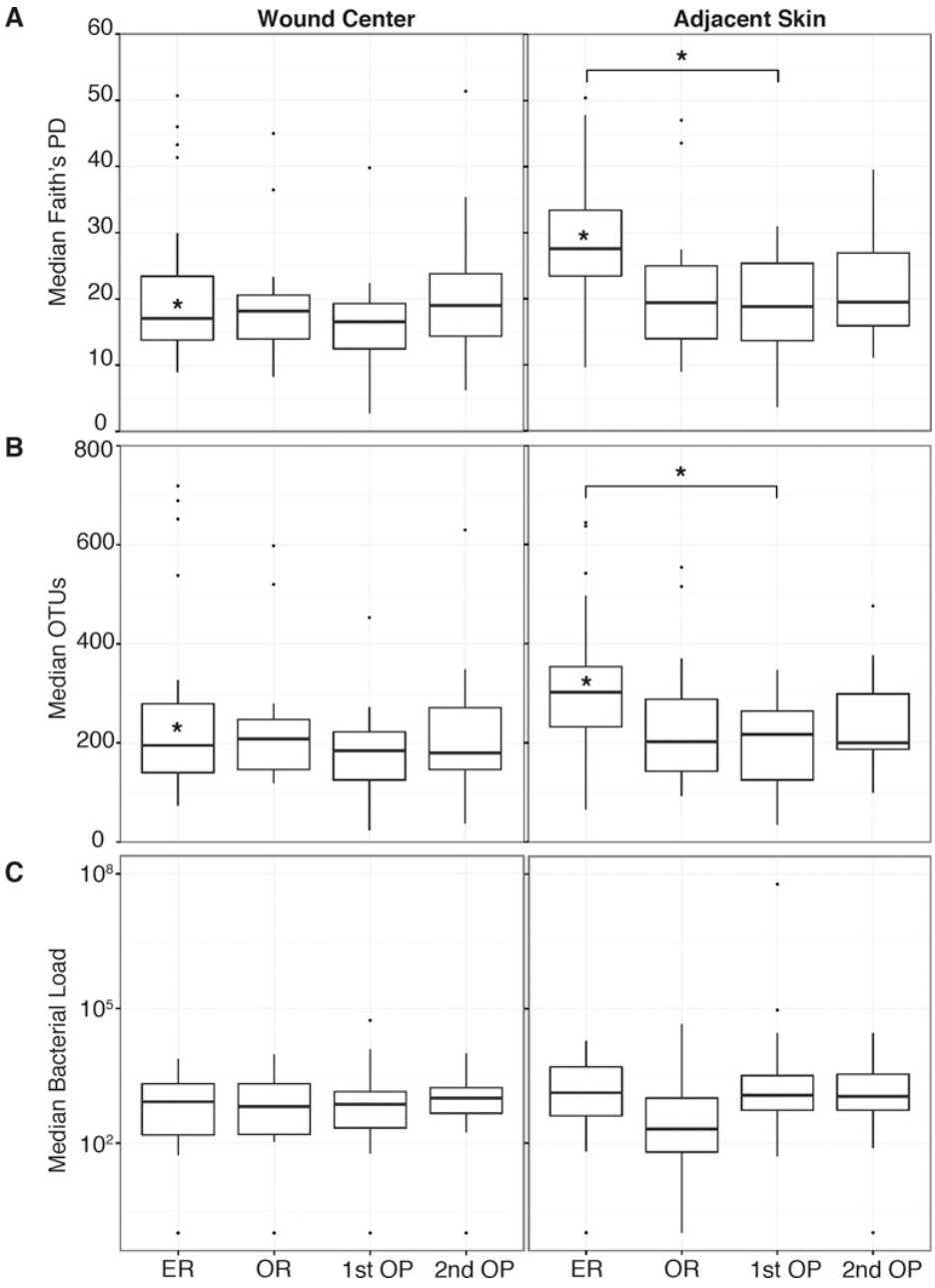

Figure 2 Alpha diversity and bacterial load of open fracture wound and adjacent skin

Alpha diversity is depicted as measured by Faith's PD (A) and observed species-level OTUs (B). Bacterial load (C) is represented as estimates from quantitative PCR of the 16S rRNA gene. The upper and lower box hinges correspond to the first and third quartiles (25% and 75%), and the distance between the first and third quartiles is defined as the inter quartile range (IQR). Lines

within the box depict median, and the whiskers extend to the highest and lowest values within 1.5 times the IQR. Outliers of the IQR are depicted with black dots above or below the whiskers. An asterisk (*) inside the box indicates significance of $P<0.05$ (Wilcoxon rank-sum test) between the adjacent skin and open fracture wound at the indicated time point. An asterisk (*) outside of the box indicates significance of $P<0.05$ (Wilcoxon rank-sum test) between the indicated time points.

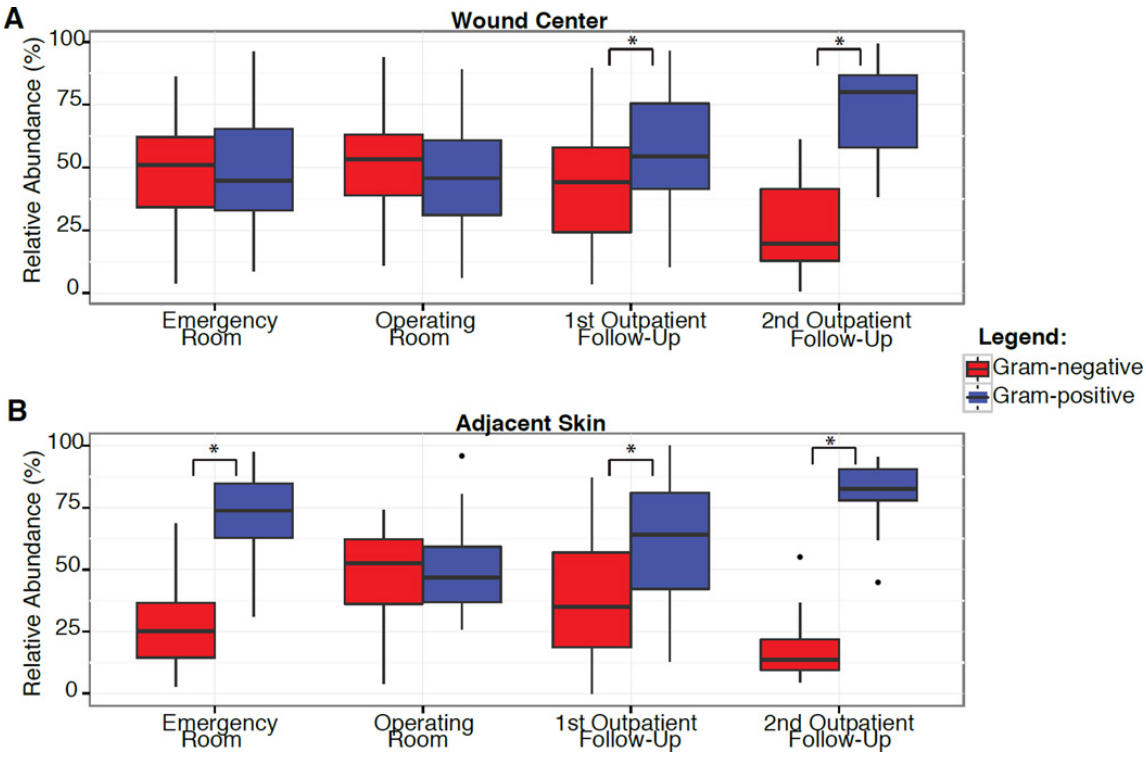

Figure 3 Gram-positive and Gram-negative bacteria in the open fracture wound and on the adjacent skin

Open fracture wound relative abundance is shown in (A) and adjacent skin relative abundance is shown in (B). The upper and lower box hinges correspond to the first and third quartiles. Lines within the box depict median, and the whiskers extend to the highest and lowest values within 1.5 times the IQR. Outliers of the IQR are depicted with black dots above or below the whiskers. *P<0.05 (Wilcoxon rank-sum test).

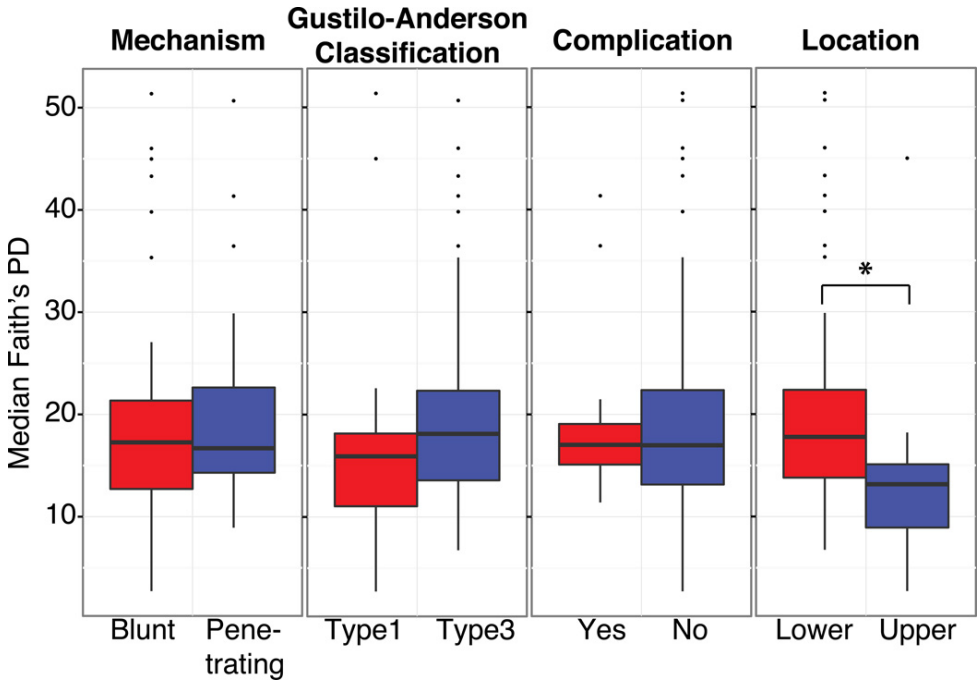

Figure 4 Association of alpha diversity with open fracture characteristics

Faith's PD comparing alpha diversity according to mechanism of injury (A), Gustilo-Anderson classification (B), whether or not the fracture healing process was complicated (C), and the anatomical location of the open fracture (D). The upper and lower box hinges correspond to the first and third quartiles. Lines within the box depict median, and the whiskers extend to the highest and lowest values within 1.5 times the IQR. Outliers of the IQR are depicted with black dots above or below the whiskers. *P<0.05 (Wilcoxon rank-sum test).

5.8 Tables

Characteristics			Cohort (n=30)
Age			
	Mean, y		43.4
	Range, y		18–82
Female Num (%)			9 (30.0)
Samples Analyzed, Wound Center			
	ER		25
	OR		13
	1st OP		21
	2nd OP		15
Samples Analyzed, Adjacent Skin			
	ER		28
	OR		12
	1st OP		21
	2nd OP		15
Mechanism (%)			
	Blunt		22 (73.3)
	Penetrating		8 (26.7)
Smoker Num (%)			
	yes		7 (23.3)
	no		20 (66.7)
	unknown		3 (10.0)
Gustilo Anderson Classification (%)			
	Type I		5 (16.7)
	Type II		1 (3.3)
	Type III		24 (80.0)
		a	22 (73.3)
		b	1 (3.3)
		c	1 (3.3)
Complication Num (%)			

Characteristics		Cohort (n=30)
	Complication	7 (23.3)
	Non-Complication	23 (76.7)
Location Num (%)		
	Tibia/fibula	12 (40.0)
	Ankle	6 (20.0)
	Femur	7 (23.3)
	Humerus	3 (10.0)
	Calcaneus	1 (3.3)
	Ulna	1 (3.3)
Hospital Stay Length		
	Mean, d	13.3
	Range, d	2–50
Days Between ER and OP1		
	Mean, d	24.3
	Range, d	13–78
	Num Samples	21
Days Between ER and OP2		
	Mean, d	47.8
	Range, d	29–69
	Num Samples	15
Antibiotics Given Num (%)		
	cefazolin	11 (36.7)
	gentamicin	2 (6.67)
	cefazolin + gentamicin	15 (50.0)
	vacomycin + gentamicin	1 (3.3)
	ampicillin + sulbactam	1 (3.3)

Table 1 Summary of cohort metadata

Genus	Wound Center				Adjacent Skin			
	ER	OR	1st OP	2nd OP	ER	OR	1st OP	2nd OP
Staphylococcus[*]	9.643	7.835	9.441	21.014	15.933	11.870	13.974	18.268
Corynebacterium	5.318[a]	2.665[b]	8.707	12.255	17.098[a]	7.454[b]	7.164	21.060
Streptococcus	3.332	2.593	3.270	4.389	2.110	4.106	4.396	3.423
Acinetobacter	3.433	1.945	4.074	3.109	1.927	2.314	1.165	1.362
Anaerococcus	0.406[c]	1.010	0.634	0.858	1.806[c]	1.610	0.826	1.130
Pseudomonas[*]	3.048[d]	2.589	0.990	1.062	0.726[d]	1.769	0.454	0.668
Unclassified Enterobacteriaceae	0.451	0.733	0.627	0.248	0.404	0.949	0.249	0.363
Propionibacterium	0.007	0.000	0.031	0.026	0.018	0.039	0.035	0.018
Escherichia[†]	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000

Table 2 Median relative abundance of taxa in open fracture wound center and adjacent skin and change over time

*Changes between time points in wound center and adjacent skin are significant ($p<0.05$; Kruskall-Wallis test)
†Changes between time points in adjacent skin are significant ($p<0.05$; Kruskall-Wallis test)
a,b,c,dDifference between skin and wound center is significantly different ($p<0.05$; Wilcoxon rank-sum test)

Factor	Taxon	Median Taxon Relative Abundance (%)		Timepoint(s)	P-value*
Mechanism	*Corynebacterium*	Blunt	5.786	1st OP	0.006
		Penetrating	30.656		
	Unclassified Enterobacteriaceae	Blunt	0.877	1st OP	0.038
		Penetrating	0.184		
	Pseudomonas	Blunt	0.944	2nd OP	0.048
		Penetrating	1.966		
Gustilo-Anderson Classification	*Acinetobacter*	Type I	4.732	All	0.015
		Type III	2.028		
	Propionibacterium	Type I	0.000	All	0.038
		Type III	0.017		
Location	*Anaerococcus*	Lower Extremity	0.858	All	0.015
		Upper Extremity	0.115		
		Lower Extremity	0.968	1st OP	0.012
		Upper Extremity	0.000		

Table 3 Taxa associated with clinical factors

*FDR-corrected P-values

5.9 Supplemental Methods

DNA Isolation

Each sample collection swab was placed in 300 µl of Yeast Cell Lysis Solution (Epicentre MasterPure Yeast DNA Purification kit) and 0.5 µl of ReadyLyse Lysozyme solution (Epicentre) was added before incubation for 1 hour at 37°C with shaking. Samples were then processed with bead beating for 10 minutes at maximum speed on a vortex mixer with 0.5 mm glass beads (MoBio), followed by a 30 minute incubation at 65°C with shaking. Subsequent steps were performed as previously described [21].

16S rRNA sequence processing and analyses

FASTQ files were generated from raw BCL files using 'configureBclToFastq.pl' (Illumina Inc.) and paired-ends were assembled using the PANDAseq Assembler [21]. QIIME 1.6.0 [24] was used for the initial stages of sequence analysis: potential sequencing artifacts outside of the 248-255 base pair length window were removed, sequences were clustered into OTUs (operational taxonomic units, a proxy for 'species') using the UCLUST method [25] at 97% sequence similarity, sequences were taxonomically classified using the RDP classifier [26] at a confidence threshold of 0.8, unclassified sequences and sequences derived from plastids were removed, samples with less than 2500 sequences were removed from the data set, each 16S amplicon pool was subsampled at an even depth of 2500 sequences for downstream processing, alpha diversity metrics (Faith's Phylogenetic Diversity index[5] and observed species-level OTUs) and a beta diversity metric (Bray-Curtis index) were calculated for each sample.

5.10 References

1. Zalavras CG, Patzakis MJ, Holtom PD, Sherman R: **Management of open fractures.** *Infect Dis Clin North Am* 2005, **19:**915-929.
2. Burns TC, Stinner DJ, Mack AW, Potter BK, Beer R, Eckel TT, Possley DR, Beltran MJ, Hayda RA, Andersen RC, et al: **Microbiology and injury characteristics in severe open tibia fractures from combat.** *Journal of Trauma and Acute Care Surgery* 2012, **72:**1062-1067.
3. Valenziano CP, Chattar-Cora D, O'Neill A, Hubli EH, Cudjoe EA: **Efficacy of primary wound cultures in long bone open extremity fractures: are they of any value?** *Arch Orthop Trauma Surg* 2002, **122:**259-261.
4. Chehoud C, Rafail S, Tyldsley AS, Seykora JT, Lambris JD, Grice EA: **Complement modulates the cutaneous microbiome and inflammatory milieu.** *Proc Natl Acad Sci U S A* 2013, **110:**15061-15066.
5. Grice EA, Segre JA: **Interaction of the microbiome with the innate immune response in chronic wounds.** *Adv Exp Med Biol* 2012, **946:**55-68.
6. Grice EA, Snitkin ES, Yockey LJ, Bermudez DM, Program NCS, Liechty KW, Segre JA: **Longitudinal shift in diabetic wound microbiota correlates with prolonged skin defense response.** *Proc Natl Acad Sci U S A* 2010, **107:**14799-14804.
7. Naik S, Bouladoux N, Wilhelm C, Molloy MJ, Salcedo R, Kastenmuller W, Deming C, Quinones M, Koo L, Conlan S, et al: **Compartmentalized control of skin immunity by resident commensals.** *Science* 2012, **337:**1115-1119.
8. Dellinger EP, Miller SD, Wertz MJ, Grypma M, Droppert B, Anderson PA: **Risk of infection after open fracture of the arm or leg.** *Arch Surg* 1988, **123:**1320-1327.
9. Johnson EN, Burns TC, Hayda RA, Hospenthal DR, Murray CK: **Infectious complications of open type III tibial fractures among combat casualties.** *Clin Infect Dis* 2007, **45:**409-415.
10. Chen AF, Schreiber VM, Washington W, Rao N, Evans AR: **What is the rate of methicillin-resistant Staphylococcus aureus and Gram-negative infections in open fractures?** *Clin Orthop Relat Res* 2013, **471:**3135-3140.
11. Murray CK, Obremskey WT, Hsu JR, Andersen RC, Calhoun JH, Clasper JC, Whitman TJ, Curry TK, Fleming ME, Wenke JC, et al: **Prevention of infections associated with combat-related extremity injuries.** *J Trauma* 2011, **71:**S235-257.
12. Murray CK, Yun HC, Griffith ME, Thompson B, Crouch HK, Monson LS, Aldous WK, Mende K, Hospenthal DR: **Recovery of multidrug-resistant bacteria from combat personnel evacuated from Iraq and Afghanistan at a single military treatment facility.** *Mil Med* 2009, **174:**598-604.
13. Wolcott R, Costerton JW, Raoult D, Cutler SJ: **The polymicrobial nature of biofilm infection.** *Clin Microbiol Infect* 2013, **19:**107-112.
14. Saltzman MD, Marecek GS, Edwards SL, Kalainov DM: **Infection after shoulder surgery.** *J Am Acad Orthop Surg* 2011, **19:**208-218.
15. Athwal GS, Sperling JW, Rispoli DM, Cofield RH: **Acute deep infection after surgical fixation of proximal humeral fractures.** *J Shoulder Elbow Surg* 2007, **16:**408-412.
16. Grice EA, Kong HH, Conlan S, Deming CB, Davis J, Young AC, Bouffard GG, Blakesley RW, Murray PR, Green ED, et al: **Topographical and temporal diversity of the human skin microbiome.** *Science* 2009, **324:**1190-1192.
17. Costerton JW: **Biofilm theory can guide the treatment of device-related orthopaedic infections.** *Clin Orthop Relat Res* 2005**:**7-11.

18. Fernandes A, Dias M: **The Microbiological Profiles of Infected Prosthetic Implants with an Emphasis on the Organisms which Form Biofilms.** *J Clin Diagn Res* 2013, **7:**219-223.
19. Schmidt AH, Swiontkowski MF: **Pathophysiology of infections after internal fixation of fractures.** *J Am Acad Orthop Surg* 2000, **8:**285-291.
20. Gustilo RB, Anderson JT: **Prevention of infection in the treatment of one thousand and twenty-five open fractures of long bones: retrospective and prospective analyses.** *J Bone Joint Surg Am* 1976, **58:**453-458.
21. Gardner S, Hillis S, Heilmann K, Segre J, Grice E: **The neuropathic diabetic foot ulcer microbiome is associated with clinical factors.** *Diabetes* 2013, **62:**923-930.
22. Hodkinson BP LF: **A microbiotic survey of lichen-associated bacteria reveals a new lineage from the Rhizobiales.** *Symbiosis* 2009**:**163–180.
23. Grice EA, Kong HH, Renaud G, Young AC, Bouffard GG, Blakesley RW, Wolfsberg TG, Turner ML, Segre JA: **A diversity profile of the human skin microbiota.** *Genome Res* 2008, **18:**1043-1050.
24. Masella AP, Bartram AK, Truszkowski JM, Brown DG, Neufeld JD: **PANDAseq: paired-end assembler for illumina sequences.** *BMC Bioinformatics* 2012, **13:**31.
25. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, et al: **QIIME allows analysis of high-throughput community sequencing data.** *Nat Methods* 2010, **7:**335-336.
26. Edgar RC: **Search and clustering orders of magnitude faster than BLAST.** *Bioinformatics* 2010, **26:**2460-2461.

CHAPTER 6 – patPRO: An R package for the visualization of longitudinal microbiome data

6.1 Abstract

6.1.1 Background

Longitudinal profiling of human microbiomes during the course of disease or perturbation provides valuable insights into disease associations and potential mechanisms that are not possible with cross-sectional study designs. Tools to longitudinally visualize multiple microbiomic metrics (i.e. alpha diversity, relative abundance) are crucial for revealing temporal patterns and associating them with patient-level metadata including clinical outcomes.

6.1.2 Results

Here we present patPRO, an R package to quickly and easily visualize longitudinal microbiome profiles using standard output from the QIIME microbiome analysis toolkit. Multiple microbiomic metrics can be visualized simultaneously and overlaid with metadata representing clinically significant events (i.e. interventions, disease states). The modular nature of patPRO allows for easy integration into other R workflows, and the package can be applied to data from other microbiome analysis toolkits, including mothur. We demonstrate the general utility of patPRO using an example dataset of the open fracture microbiome during healing.

6.1.3 Conclusions

We show that patPRO is a user-friendly tool that allows rapid visualization of longitudinal microbiome data. This package can be applied to any longitudinal microbiome study to facilitate comprehensive summaries of single or multiple subjects, without an extensive need of R programming skills. PatPRO is optimized to work with established microbiome software

including Qiime and mothur. PatPRO will freely available on CRAN and runs on Mac OSX and Linux.

6.2 Background

Microbiome community profiling (i.e. 16S ribosomal RNA sequencing) is an increasingly accessible experimental approach, due in large part to decreasing DNA sequencing costs, increasing throughput, and greater accessibility to bioinformatics tools. Longitudinal applications of microbiome community profiling have revealed important associations between disease states and the microbiome, including communities of the skin [1], oral cavity [2], vagina [3], and gut [4]. Temporal microbiome variation is often significant across all body sites, and should also be taken into account when linking health or disease states with microbial communities [5]. Current tools to longitudinally and simultaneously visualize multiple metrics of the microbiome (i.e. alpha diversity, taxonomy) overlaid with clinically significant events are lacking.

Longitudinal visualization of microbiome data over time allows both researchers and clinicians to easily assess trends in a patient's individual microbial community, including the effects of clinically significant events such as antibiotic administration or surgical intervention. While the QIIME toolkit and R packages including qiimr and vegan provide valuable microbiome analysis and visualization tools, they do not provide straightforward, automated profiling of longitudinal data overlaid with subject metadata [6-8]. To address this, we developed the patPRO microbiome analysis and visualization package for the R statistical language, a widely used and freely available statistical software environment [9].

6.3 Implementation

PatPRO is an R package that facilitates rapid visualization of longitudinal microbiome data. This package implements data munging and plotting functionality as a user-friendly toolkit that requires minimal R programming experience. The patPRO package was optimized for QIIME-formatted files [8], but can also be used with other analysis toolkits, such as mothur [10]. PatPRO's modular design for data formatting, calculations, and plotting allows for flexibility in workflow design, customizability of the output, and it can be used in combination with other packages or analysis workflows.

6.4 Results & Discussion

patPRO uses four input files to generate patient profiles, as outlined in **Figure 1A**: a QIIME-formatted mapping file that associates the sample identifiers with their metadata, a QIIME taxonomic relative abundance table, a QIIME alpha diversity table, and an optional bacterial load data table, which is obtained by 16S quantitative PCR (qPCR) analysis.

6.4.1 Initial Data Formatting

Using the *merge.map.meta.data* function of the patPRO package, each input table is merged with the metadata outlined in the mapping file. Merging is required for further downstream processing.

6.4.2 Taxonomic Profiling

Once merged, the taxonomic relative abundance table is used to determine the most abundant taxa within the patient of interest using the top.rel.abund.data.frame function. This function automatically formats the data table for visualization with the plot.top.taxa function (Figure 1B; bottom panel). Absolute bacterial abundance information (also called bacterial load) from qPCR

experiments is visualized directly using the plot.bacterial.load function (Figure 1B; middle panel), or used to normalize the relative abundance profiles using the top.abs.abund.data.frame function (Figure 1C; bottom panel). The normalized relative abundance profiles are visualized using the top.abs.abund.plot function. The normalization process calculates the taxonomic relative abundance as a percent of the overall bacterial abundance. Important events are added to both standard and normalized relative abundance plots using the arrow annotation functionality.

6.4.3 Alpha Diversity Profiling

The normalize.alpha.div function normalizes the desired alpha diversity metric values to a percent of the first value (i.e. the first time point). This normalization is necessary because different alpha diversity metrics have orders of magnitude differences in scale, making plotting of different un-normalized metrics on the same graph impractical. This function also automatically formats the data table for visualization using the plot.normalized.alpha.div function (Figure 1B; middle panel).

6.4.4 Average Patient Profiling

In addition to processing and visualizing individual patient profiles, patPRO offers tools to analyze the averages of multiple patients. The patient.mean function, in addition to the other related patPRO functions, allows for easy calculations and visualizations of the mean and standard error of taxonomic relative abundance, diversity, and bacterial load values (Figure 1D).

6.4.5 Comprehensive Patient Profile Visualization

Using the individual microbiome plots above, patPRO generates single, comprehensive profiles containing two or three microbiome plots, using the patpro.plot.two and patpro.plot.three functions, respectively (Figure 1B-D).

6.4.6 patPRO Analysis Example

To illustrate the use of patPRO, we provide example data from a longitudinal microbiome study of open fractures from patient presentation to the emergency department, through surgical intervention and outpatient follow up (**Figure 1B-D**) [11]. Using patPRO, we observed shifts in microbial diversity, bacterial load, and relative abundance associated with surgical interventions (**Figure 1B-C**). Following the second operation, we observed an increase in relative and absolute abundance of Staphylococcaceae bacteria (**Figure 1C**). This example dataset is included in the patPRO package and an analysis workflow example is included as an R notebook accompanying this manuscript (Additional File 1).

6.5 Conclusions

PatPRO is a user-friendly tool that allows rapid visualization of longitudinal microbiome data, facilitating exploratory analyses of patient-specific trends and association of microbiome with disease states, perturbations, and/or clinical outcomes. We believe patPRO will have an especially large impact for clinical studies because changes in the microbiome can be tracked for long time periods and associated with patient-level metadata and/or clinical outcomes. PatPRO is available under a GNU General Public License (GPLv3) and will soon be available for public use on CRAN.

6.6 Acknowledgements

Funding: This work was supported by grants from the National Institutes of Health (NIAMS R00 AR060873 to EAG), the Pennsylvania Department of Health, and the Orthopaedic Trauma Association. GDH was supported by the US Department of Defense (DoD) National Defense Science and Engineering Graduate (NDSEG) Fellowship. MAL and BPH were supported by a grant to the University of Pennsylvania Department of Dermatology (NIAMS T32 AR007465).

6.7 Figures

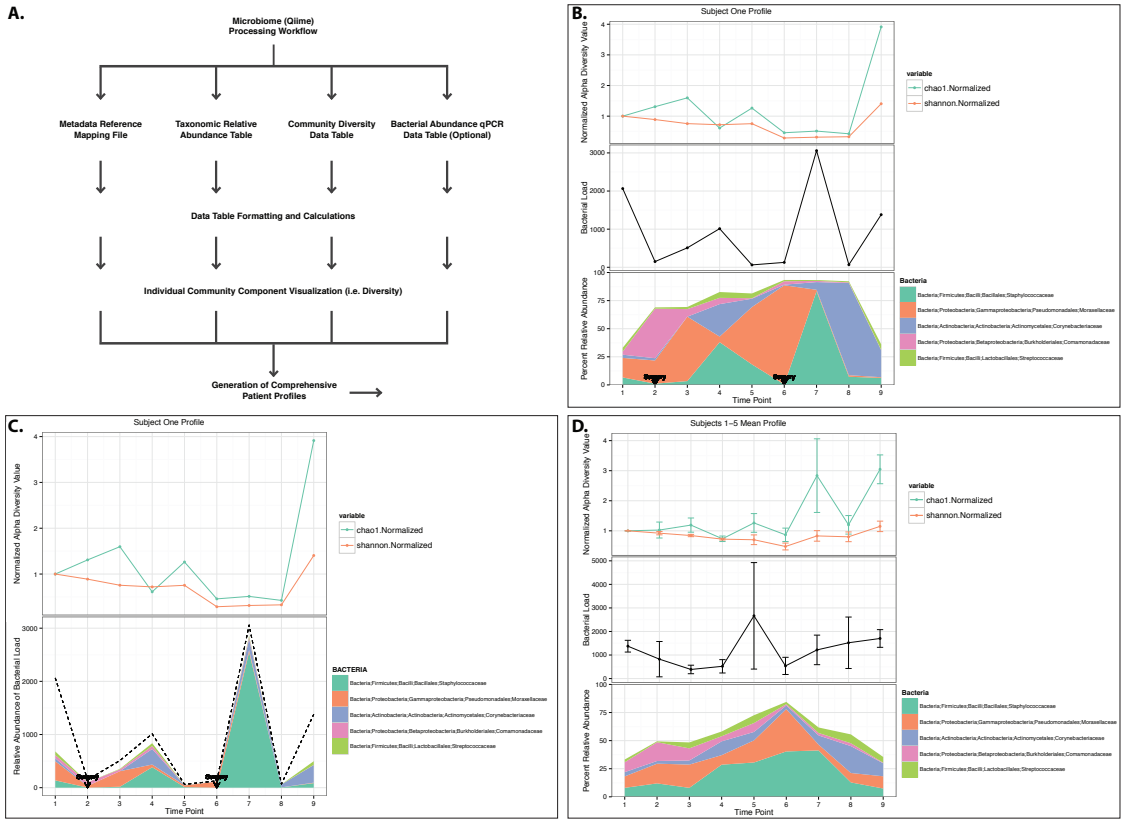

Figure 1 PatPRO provides multiple comprehensive plotting options for longitudinal microbiome data

A) Flow chart outlining the data processing and visualization steps included in the patPRO package. B) Example of a single patient profile including normalized alpha diversity (top panel), bacterial load (middle panel), and relative abundance with annotated surgical events (bottom panel). C) Single patient profile including taxonomic relative abundance normalized to bacterial load (bottom panel). D) Example of a patient profile depicting the mean values of five patients, compared to a single patient profile in section (B). Bars represent standard error.

6.8 References

1. Kong HH, Oh J, Deming C, Conlan S, Grice EA, Beatson MA, Nomicos E, Polley EC, Komarow HD, Murray PR, et al: **Temporal shifts in the skin microbiome associated with disease flares and treatment in children with atopic dermatitis.** *Genome Res* 2012, **22:**850-859.
2. Huang S, Li R, Zeng X, He T, Zhao H, Chang A, Bo C, Chen J, Yang F, Knight R, et al: **Predictive modeling of gingivitis severity and susceptibility via oral microbiota.** *ISME J* 2014, **8:**1768-1780.
3. Gajer P, Brotman RM, Bai G, Sakamoto J, Schutte UM, Zhong X, Koenig SS, Fu L, Ma ZS, Zhou X, et al: **Temporal dynamics of the human vaginal microbiota.** *Sci Transl Med* 2012, **4:**132ra152.
4. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen YY, Keilbaugh SA, Bewtra M, Knights D, Walters WA, Knight R, et al: **Linking long-term dietary patterns with gut microbial enterotypes.** *Science* 2011, **334:**105-108.
5. Flores GE, Caporaso JG, Henley JB, Rideout JR, Domogala D, Chase J, Leff JW, Vazquez-Baeza Y, Gonzalez A, Knight R, et al: **Temporal variability is a personalized feature of the human microbiome.** *Genome Biol* 2014, **15:**531.
6. Bittinger K: **qiimer: Work with QIIME Output Files in R.** 2014.
7. Jari Oksanen FGB, Roeland Kindt, Pierre Legendre, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens and Helene Wagner: **vegan: Community Ecology Package.** 2015.
8. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, et al: **QIIME allows analysis of high-throughput community sequencing data.** *Nat Methods* 2010, **7:**335-336.
9. Team RC: **R: A language and environment for statistical computing.** *R Foundation for Statistical Computing, Vienna, Austria* 2014.
10. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, et al: **Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities.** *Appl Environ Microbiol* 2009, **75:**7537-7541.
11. Hannigan GD, Hodkinson BP, McGinnis K, Tyldsley AS, Anari JB, Horan AD, Grice EA, Mehta S: **Culture-independent pilot study of microbiota colonizing open fractures and association with severity, mechanism, location, and complication from presentation to early outpatient follow-up.** *J Orthop Res* 2014, **32:**597-605.

Together these studies represent a significant step forward in understanding the skin microbiome in healthy and diseased states. In healthy skin, we presented evidence that the virus communities are primarily diverse phage populations, their community composition depends on anatomical locations, and that they have a strong potential for promoting horizontal gene transfer by transduction, especially of antibiotic resistance genes and virulence factors. We further suggested that the phage communities have complex interactions with their bacterial hosts, especially in the context of CRISPR targeting. These findings highlight the importance of micro-environmental context when considering dermatological viruses. Viruses studied at one location are in a different community than those studied at a different location, and this may impact virus behavior. We also highlight the potential of the skin virome for persisting as latent prophages and facilitating horizontal gene transfer of clinically relevant genes, including those conferring antibiotic resistance. These novel insights into the virome present a foundation for future studies of the skin virome in diseased or perturbed states. These findings also lay a foundation for studies of the associations between the skin virome and the environment, especially related to transduction of pathogenic bacterial genes.

As our primary future direction for this study, we will be using the findings as the foundation for understanding the genomic variability associated with the skin virome. Our high virus genomic coverage provides us with a unique opportunity to study the evolutionary and functional implications of single nucleotide polymorphisms within the natural skin virome. We will use this approach to provide an understanding of virus genomic variability and evolution in a natural skin community, which to date has not been studied.

While studying the healthy skin virome, we also investigated the microbial communities of the understudied, at-risk acute open fracture wounds. This represents a new skin disruption state whose microbiome has not yet been robustly evaluated. We found that the wound microbiome becomes more similar to the adjacent skin (healthy) microbiome as healing progresses. This suggests a role for microbiome healing along with tissue healing in acute wounds. We also noted numerous associations between the microbiome factors and important clinical factors, including severity and the progression to complications. We finally presented a tool that was used for processing and visualizing the longitudinal microbiome dataset, and will be publicly available for other researchers to use in a similar capacity.

Because the presented study was only a pilot with half of the cohort recruited, our obvious next step will be evaluating the open fracture microbiome using the complete cohort. The complete cohort dataset will include twice as many patients with follow up data up to one year after presentation to the Emergency Room. Analysis of this extended follow up timeline will provide us with insight into the long-term dynamics of the microbiome during open fracture healing. This is in comparison to the relatively short timeline (~50 days) in the presented pilot study.

The increased sample size will allow us to better evaluate the associations between the microbiome and clinical complications, including infections. The larger cohort will also offer an opportunity to evaluate the utility of the initial wound microbiome in predicting the outcome of the open fracture healing. This will be done using predictive modeling approaches that will allow us to predict the wound healing outcome based on microbiome signatures from when the patient

presents to the emergency room, including microbial relative abundance and diversity. This is an exciting avenue for this research program because it benefits the patient by allowing us to identify those wounds at risk for microbial complications, and it is a non-invasive method that can relatively easily be implemented in the clinic.