# Dominant Effect Of Host Genetics On Skin Microbiota Composition In Homeostasis And Wound Healing

Jack Galbraith[1], Julien M. D. Legrand[1,2], Nicholas Muller[1,3], Betoul Baz[1,4], Katie Togher[3], Nicholas Matigian[5], Seungha Kang[3], Sylvia Young[7], Sally Mortlock[6], Edwige Roy[1], Grant Morahan[7], Graeme Walker[1]*, Mark Morrison[3]*, Kiarash Khosrotehrani[1]*

1. The University of Queensland, UQ Diamantina Institute, Experimental Dermatology Group, Brisbane, Australia

2. Australian Regenerative Medicine Institute, Monash University, Clayton, Victoria, Australia.

3. The University of Queensland, UQ Diamantina Institute, Microbial biology and metagenomics group, Brisbane, Australia

4. National Centre for Genomic Technologies, King Abdul Aziz City for Science and Technology, Riyadh, Saudi Arabia

5. QCIF Facility for Advanced Bioinformatics, Institute for Molecular Bioscience, The University of Queensland, St Lucia, Brisbane, QLD Australia

6. Institute for Molecular Bioscience, The University of Queensland, St Lucia, Brisbane, QLD Australia

7. Centre for Diabetes Research, The Harry Perkins Institute for Medical Research, University of Western Australia, Perth, WA, Australia

*Contributed equally as senior authors

Corresponding author:

Kiarash Khosrotehrani, MD PhD FACD

The University of Queensland Diamantina Institute

Translational Research Institute

37 Kent Street, Woolloongabba 4102 QLD, Australia

Email: k.khosrotehrani@uq.edu.au

## Abstract

Animal microbiota are shaped and maintained not only through microbiota-environmental interactions but also through host-microbiota interactions. The effects of the microbiota on the host has been the source of intense research in recent years, indicating a role for resident microbes in a range of conditions from obesity and mood disorders to atopic dermatitis and chronic wounds. Yet the ability of hosts to determine their microbiota composition is less well studied. In this study, we investigated the role host genetics plays in determining skin microbiota. We used 30 different mouse strains from the advanced recombinant inbred mouse panel, the Collaborative Cross, with PERMANOVA, GWAS and PCA-based GWAS analyses to demonstrate that murine skin microbiota composition is strongly dependent on murine strain. In particular, a quantitative trait locus on chromosome 4 associates both with *Staphylococcus* abundance and principal-component multi-trait analyses. Additionally, we used a full thickness excisional wound healing model to investigate the relative contributions from the skin microbiota and/or host genetics on wound healing speed. Wound associated changes in skin microbiota composition were observed and were in many instances host-specific. Despite reaching statistical significance, the wound-associated changes in skin microbiota accounted for only a small amount of the variance in wound healing speeds, with the majority attributable to mouse genotype (strain) and age. Host genetics has a significant impact on the skin microbiota composition during both homeostasis and wound healing. These findings have long reaching implications in our understanding of associations between microbiota dysbiosis and disease.

## Introduction

All animals are colonised by microbes soon after birth. In recent years, these "microbiomes" have been implicated in a wide range of host responses relevant to homeostasis, and their disruption can manifest in a wide range of immune-mediated and/or metabolic-related diseases [1, 2]. The gut (stool) microbiome has been the most intensively studied and has advanced the concept of "community-based" interactions, that trigger a range of conditions such as atopic dermatitis, obesity or mood disorders [3-5]. Although less studied, variations in the skin microbiome have been associated with episodes of atopic dermatitis [6], and the healing rate of leg ulcers, among other conditions [7, 8]. The skin microbiome has been shown to vary with anatomical location and patient age [9], and environmental cues such as humidity and/or temperature may also influence inter-individual variation [10]. In that context, diet is now also considered to exert a strong effect on the microbiome that can in turn influence host response and health status [2]. Overall, many environmental factors seem to determine the composition of the skin microbiota.

However, there are relatively few studies that provide a systematic assessment of how host genotype affects microbiome composition. Broader genome-wide studies utilising inter-cross mouse panel (BXD) have identified many host-specific quantitative trait loci (QTL) affecting gut microbiome composition, with candidate genes involving cytokines and toll-like receptor signalling [11]. In addition, investigation of the gut microbiota of mice from another advanced inter-cross model found 18 separate QTL associated with the relative abundance of one or more gut bacterial taxonomies [12]. Regarding skin, most studies to date have been limited to candidate approaches, such as how mutations affecting the serine protease matriptase can lead to a shift in the skin microbiota composition [13], or have described the skin microbial populations more generally [14-16]. In addition, many studies investigating wounds and skin microbiome associations have focused on non-healing chronic wounds such as diabetic foot ulcers and venous leg ulcers [17-19]. Whether host genetics also affects the wound microbiome, or alternatively, whether the environmental changes associated with wounding dominate over the effect of host genetic and more strongly affect microbial composition, are unknown.

Here we utilised a resource generated by a specialised breeding program, the Collaborative Cross (CC), which through a recombinant-inbreeding design, allows interrogation of complex traits and genetic pleiotropies [20, 21]. Using 30 different CC mouse

strains, we investigated the host genetic contribution to mouse dorsal skin microbiota composition during homeostasis and wound healing. We found that the cutaneous microbiota composition differed between mouse strains, and that responses to wounding related to their microbiome composition were strain-specific. Genome-wide association studies (GWAS) identified key QTL associated with specific bacterial taxa from normal skin. Host genetics not only accounts for most of the observed variation in microbiome composition of normal skin, but also affects wound healing speed, while microbiota composition was found to only have a limited role in the latter process.

## Results

### *Murine strain is a strong determinant of microbiota diversity*

The diversity of the skin microbiome has been shown to vary between individuals and can often associate with pathological states. Here we collected swabs of mouse dorsal skin for 16S rRNA gene amplicon sequencing to investigate differences in microbiome composition between mouse strains. For each mouse, swabs were collected in a predefined area of the dorsal skin in a reproducible way. Control swabs (sham swabbing), swabs taken of skin immediately after full-excisional wounding, for each mouse, did not result in any significant amplification of 16S rRNA gene. We first compared the Shannon's diversity metric as a measure of alpha-diversity (within sample) across 114 mice from 30 different mouse strains using the R package 'vegan'. The microbiota profiles from most mice fell within a relatively narrow range of Shannon's diversity values (IQR=0.59, median=2.11), but communities in some mice clearly displayed less diversity and/or evenness (**Fig. 1a**). Additionally, the intra-strain variation in Shannon's diversity values was generally small, but the inter-strain variation, as measured by Kruskal-Wallis testing was large and significant ($p<0.0001$). Importantly, the clustering of mice within their respective strains was also established, more generally, with their skin microbial composition (family level) by Ward's method of agglomerative hierarchical clustering [22, 23] (**Fig. 1b**). This analysis revealed that the large heterogeneity in overall mouse skin microbiota composition was strongly strain-dependent, which was further confirmed by PERMANOVA (R-squared > 0.5, $p=0.001$). We next asked whether specific microbial genera were associated specifically with lower or higher diversity compositions ('ALDEx2' R package). The abundances of both *Staphylococcus* and *Aerococcus* spp. significantly differed across the highest and lowest quartiles of Shannon's diversity values,

respectively (*Staphylococcus* $p$=0.0019, *Aerococcus* $p$=0.0028). A linear regression model using Shannon's diversity and the centred-log ratio (CLR) transformation of the abundance values for *Staphylococcus* and *Aerococcus* across all mice showed there was an inverse relationship of both genera with the Shannon's diversity value of mouse skin microbiome (**Supp. 2c**).

***Skin microbiota composition is determined by murine strain specific genomic loci.***

Next, we determined a core skin microbiome and used different levels of bacterial taxonomies to identify their representation in at least 50% of samples at 0.1% or greater relative abundance. Similar to the broader composition, core microbiome of individual mice clustered within their respective strains (Ward's method of agglomerative hierarchical clustering) and significant groupings at the family-level of classification were found and could be further used to categorise the different host genotypes (**Fig. 1b)**. While *Staphylococcaceae* was the most prevalent bacterial family in the strains of mice examined, and thereby a member of the core microbiome, there was also a group of mouse strains that clustered based on their possession of a relatively high abundance of *Corynebacteriaceae* (**Fig. 1b & Fig. 1c**). Although mouse age significantly affected skin microbiome profile, it was found to explain only a minimal amount of the variance observed across mice (R-squared < 0.02). Taken together, these findings strongly confirmed the effect of murine strain (host genotype) as a key determinant of the dorsal skin microbiome.

Given these findings, we next examined whether specific loci in the mouse genome could be associated with the dorsal skin microbiome composition. The centred log-ratio transformation of the abundance values for *Staphylococcus* was used as a "trait" for host genome-wide association analysis using the GeneMiner software, and identified a genome-wide significant region on mouse chromosome (Chr) 4 between 129.75-130.95 megabase pairs (Mbps) (LOD Score: 10, **Supp. 1a**). Several different loci, albeit with weaker (suggestive) LOD scores were identified using both centred-log ratio transformation of the abundance values for *Aerococcus* (Chr 13, 108.70-113.51 Mbps, LOD Score: 6, **Supp. 1b**), and Shannon's diversity scores (Chr 15, 3.20-7.40 Mbps, LOD Score: 8, **Supp. 1c**), along with principal component analysis (PCA)-based GWAS, which is an approach to identify multi-trait loci from multi-dimensional data [24] (**Supp. 1d-e**). Interestingly, the peak on Chr 4 identified for *Staphylococcus* was also recovered using the PCA-based GWAS (**Supp. 1e**). A detailed analysis of the founder haplotypes for the Chr 4 candidate region identified the WSB and PWK

founder alleles associated with a relatively low abundance of *Staphylococcus* in skin microbiome, and the CAST founder allele with highest relative abundances. We next examined this region of interest for any genes harbouring specific polymorphisms in the founder haplotypes above, by using the Sanger Mouse Genome Project SNP query website. Here, we identified *Ptafr* as a candidate gene of interest, on the basis that its product may affect the immune response to pathogens [25, 26], but is also known to act directly on the wound healing process [27], and can affect skin inflammation [28]. An alternative candidate is *Smpdl3b*, *Sphingomyelin Phosphodiesterase Acid-Like 3B*, which is associated with inflammation via negative regulation of toll-like receptor signalling *in vivo* [29]. A full list of genes containing haplotype specific SNP is provided in **Suppl. 3**. Overall, this strong effect of murine strain, as well as the association with plausible candidate genes, strongly supports the importance of host genetics in determining the skin microbiome during homeostasis.

***The microbiomes of early stage wounded and normal skin retain host-strain specificity***

We next investigated whether and how wounding, considered a major environmental stress, would elicit stronger effects on the microbiome than host genotype. If skin injury and wound healing had strong effects on the microbiome, one would expect the wound microbiome to be convergent across the different mouse strains. Alternatively, given the exposure to faecal material in the cage bedding, one would expect the wound and faecal microbiome to converge. To that end, we performed PCA and hierarchical clustering on the combined datasets from unwounded skin, day (D) 3 wounds, and the faecal microbiome samples of 70 mice representing the different haplotypes. Importantly, the DNA extractions of the swab samples collected immediately after wounding did not produce sufficient 16S rRNA gene amplicons, suggesting that the debrided area was made “sterile” by the wounding process at D0. We found that the wound microbiome at D3 retained its similarity to the community present on unwounded skin, and that the faecal microbiomes were clearly separable from the skin and wound microbiomes using PCA and hierarchical clustering **(Fig. 2a-b)**. The bacterial families *Propionibacteriaceae* and *Staphylococcaceae* were discriminatory for both the unwounded skin and wound microbiomes at D3, compared to faecal samples, and *Bacteroidales* family S24-7 were in much greater abundance in faecal samples (**Supp. 2a**). However, while the D3 wound microbiomes still retained features of normal skin, there was a significant decline in diversity between unwounded and D3 skin wound microbiomes (**Fig. 2c**, Mann-Whitney, *p*-value<0.0001, 0.788 median fold-change in Shannon’s diversity across all mice)**.** This decline

in diversity recapitulates that seen in other pathologies such as eczema flares, and some non-healing diabetic foot ulcers [14, 17]. The declines in microbial diversity of the wounds were only transient however, and diversity had recovered significantly by D10 post-wounding via Kruskal-Wallis testing ($p<0.01$, 1.1 median fold-change in Shannon's diversity across all mice), but still remained significantly lower than the original unwounded skin at this time-point ($p<0.01$, 0.8 median fold-change unwounded to D10 Shannon's diversity).

We also evaluated whether the 'core' microbiome differed between wounded (D3) and unwounded skin. The hierarchical clustering (**Fig. 3a**) showed that although there was a general trend of increased *Staphylococcaceae* and *Corynebacteriaceae* relative abundances between D3-wounded and unwounded skin (FDR-adjusted, $p<0.1$, ALDEX2), some mouse strains showed a decrease in these families. This clearly highlighted that the changes of microbiota composition in response to wound healing were not universal and varied across murine strains.

To investigate the relative contributions to changes in microbiota composition during wound healing, a PERMANOVA analysis was performed with the variables being mouse strain, wound time-point (unwounded or D3) and mouse age (days). The strain of mouse had the largest effect on overall microbiome compositional changes during wound healing (~23% variance explained), followed by the interaction between mouse strain and wound time-points (~17% variance explained). While statistically significant, age and wound time-point alone accounted for only a small amount of variance (<8% combined). This once again strongly supports the effect of host genotype on both the unwounded and wounded skin microbiomes.

### ***Microbiome composition had minimal effect on wound healing rates***

Given the variation between mice in their D3 post-wounding microbial composition and the microbiota differences from unwounded skin, we investigated whether these parameters were associated with wound healing. The time to wound closure for individual mice was determined and used to group these into quartiles representing fast to slow healers. Compared to unwounded skin, there were large variations, both positively and negatively, in the relative abundances of different bacterial families present in D3 wound microbiomes assigned to the different quartiles (**Fig. 3b**). However, these changes in the skin/wound microbiomes during early wound healing, and the faecal microbiome composition of the host animal each accounted for ~5% of the variance in wound healing speeds when modelled separately by principal component regression. In contrast, PEMANOVA analysis showed mouse strain

(genotype) accounted for over 50% of the variance in wound healing rate ($p<0.001$) while the age of mice explained 20% variance ($p$-value<0.001). Age is already recognised to be a significant factor in the regenerative ability of mice [30, 31]. Our results not only confirm the association between mouse age and wound healing speed, but the limited impact of age on skin microbiome composition suggests the impacts of age on wound healing speed are host- rather than microbiome-related.

## Discussion

The skin microbiome has been the subject of many studies because of its presumed involvement with the onset and/or progression of many skin disorders [5-7, 14, 17, 19, 32]. Many studies have revealed how (micro) environmental factors such as temperature, humidity, dryness, sun exposure, body site, or host factors such as age and diet can affect the skin microbiome [5, 9, 10, 33]. These collective findings have led many to infer that wounding - which is an extreme form of environmental insult - results in a stereotypical change in the composition of the microbiome at the wound site, principally via increases in the relative and total abundances of *Staphylococcus* and Gram-negative bacteria [32, 34]. In contrast, very little is known about whether and how host genotype affects the skin microbiome, and further, to what extent differences in wound healing speed are attributable to host-driven processes either directly, or indirectly, via the skin microbiome.

By characterization of multiple strains of mice from the Collaborative Cross, we show that while the skin microbiome is strongly impacted by host genotype, there is only a small contribution from this microbiome variation to wound healing speed. Further, we showed a strong variation in skin microbiota composition that in large part was explained by the murine genetic background. Importantly, all the animals used in this study were born in the same animal facility; then all shipped and subsequently housed in a different facility and provided access to the same food and water sources. So while it is not possible to completely rule out any housing effects across all 30 strains of mice, other studies have revealed that compared to host genotype, the contributions from caging and legacy effects on the variations in the gut microbiome are small [35]. Additional studies of skin bacterial populations in various mammals and amphibians further supports host taxonomy as a greater determining factor of the cutaneous microbiome than environment [36-38].

In that context, the genome-wide associations identified key host loci predictive of specific skin bacterial taxa and/or microbiome composition. Indeed, specific bacterial genera (*Staphylococcus* and *Aerococcus*) and Shannon diversity scores of the skin microbiome from individual mice could be used as a "trait" in genome wide association and linkage studies. In particular, these analyses defined a ~1 Mbp region on Chr 4 to be strongly associated with these traits. Further analyses confirmed this locus includes genes affecting host innate and adaptive immunity and in particular *Ptafr*, which encodes the Platelet Activating Factor Receptor. The PTAFR protein has strong pro-inflammatory effects and has been previously associated with bacteraemia [25, 26, 28]. Similarly, *Smpdl3b*, Sphingomyelin Phosphodiesterase Acid-Like 3B, is associated with inflammation via negative regulation of toll-like receptor signalling in vivo [29]. While the precise role of these candidate genes needs experimental validation, it is plausible that SNP variations in one or both of these genes can elicit differential immune or dermal niche alterations that affect the relative abundance of *Staphylococcus* and/or *Aerococcus* on skin.

Abundance of *Staphylococcus* and *Aerococcus* was associated with lower diversity in homeostatic skin microbiota. The *Staphylococcus* genus contains both commensal and pathogenic species, such as *staphylococcus epidermidis* and *staphylococcus aureus*. Host *NOD2* receptor variants have been associated with atopic dermatitis and can affect keratinocyte susceptibility to *S. aureus* [39]. Additionally, previous authors have shown an expansion of the *S. aureus* population is associated with pathogenic states, such as eczema flares [14]. One study in diabetic men showed a higher Staphylococcal abundance in healthy controls compared to diabetics, yet the proportion of *S. aureus* was lower in controls indicating that some less virulent Staphylococcal species are likely protective [18]. Interestingly, *S. epidermidis* can help co-ordinate the host's response to *S. aureus* by inducing host cytokines that influence T-cell behaviour [40]. Whilst our data supports other's findings of a host effect on Staphylococcal species abundance and skin microbiota diversity, it is unclear whether the changes we see in Staphylococcal abundance include expansions of pathogenic or commensal species, as we were unable to investigate taxonomies to the species level. An important implication of our study is that microbiome dysbiosis in disease may only be restored temporarily through intervention such as probiotics, as the genetic background of the individual may have a dominant role.

In a clinical setting, it has been thought that skin wounds are characterised by major colonisation of Gram-negative bacteria, such as *Pseudomonas aeruginosa* [34], and often display an over-representation of Staphylococci [32]. Here, we report that wound microbiome

remains highly variable in its composition across different murine strains and there was no homogenous change across all mice. Inclusion of stool samples allowed us to show that D3 post-wound skin retains a skin-like microbiota phenotype. This is unexpected given the extreme barrier function dysregulation as well as the relative contamination of cage bedding with faeces. Considering the site specificity and stability of the human microbiome [9], it is possible that the transcriptomic environment of skin wounds drives the microbiota to retain its skin-like features.

Few studies of murine microbiota composition of skin have utilised multiple strains of mice; a majority used only a single mouse genetic background [37, 41]. Our study highlights the difficulty of drawing conclusions about microbial associations with wound healing outcomes across studies that have used a single background mouse model of microbiota changes. Studies of chronic wounds in patients show relatively little difference in the more abundant genera between healing and non-healing wounds [42]. In line with these results, we were unable to identify any statistically significant, or suggestive, abundant microbial families associated with healing speed in mice, although we were able to show suggestive differences in microbial compositional changes between mouse strains during the early stages of wound healing. Principal component regression showed that both faecal and skin microbiomes account for minimal differences in wound healing speed. Additionally, PERMANOVA analysis shows that the strain of mouse, and its interaction with the wound time-course, can explain ~40% of variance in microbiota composition changes during the first 3 days of wound healing. Whilst lack of significant associations between microbiota abundances and wound healing speeds could be due to omission of low abundance microbes, significant heterogeneity in microbiota composition between mouse strains, and their specific response to wounding, likely plays a dominating role.

In conclusion, we report that skin murine microbiota and its changes upon wounding are strongly determined by host genetics and the abundance of specific microbial families can be determined by precise loci in the murine genome. Moreover, the wound microbiome plays a minimal role in the healing rate and is mostly a reflection of the host genetic background. These findings have far reaching implications for the design of further studies on the role of the microbiome on health and disease as well as the use of probiotics in a clinical setting.

**Figure 1.**

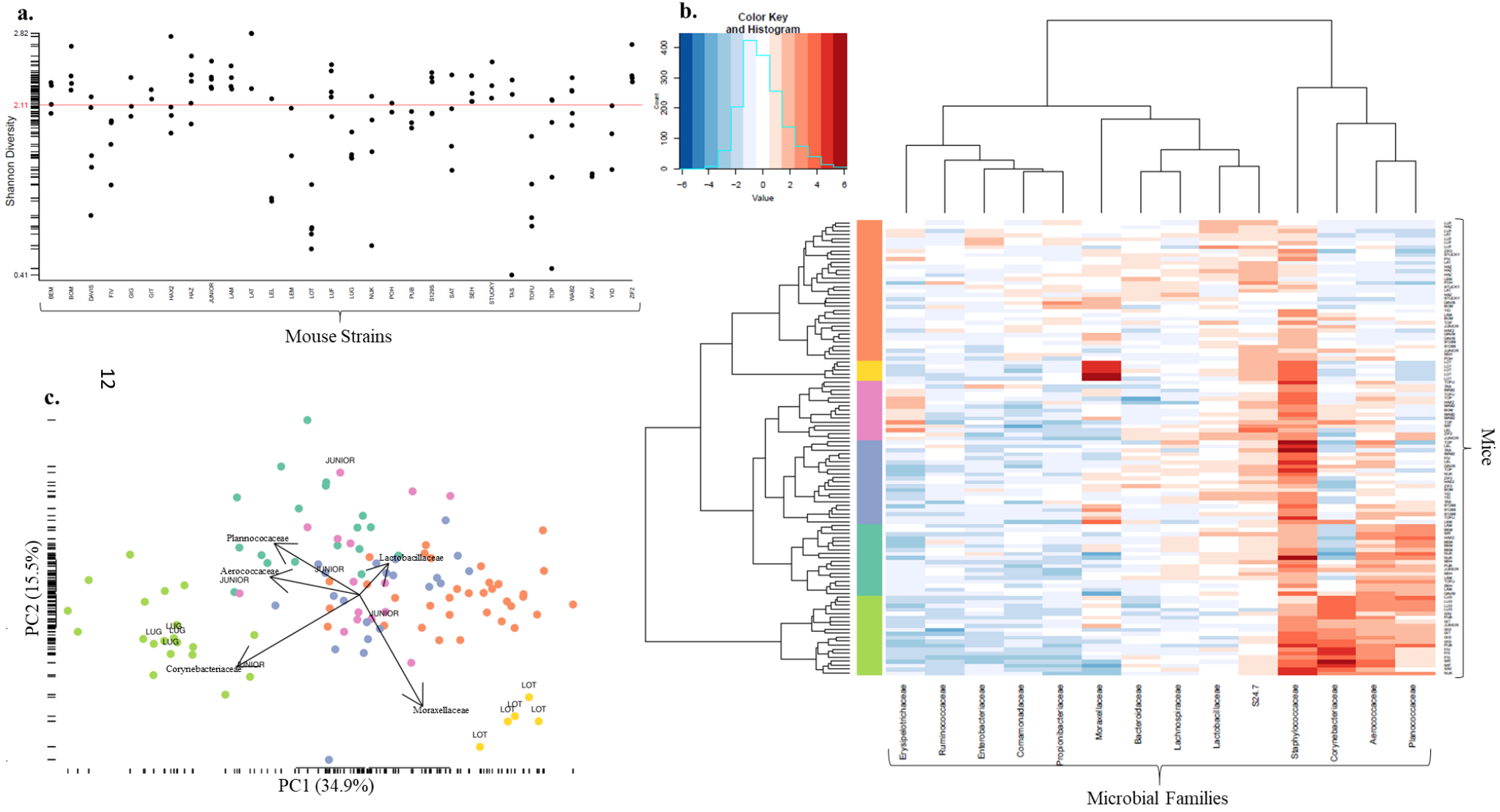
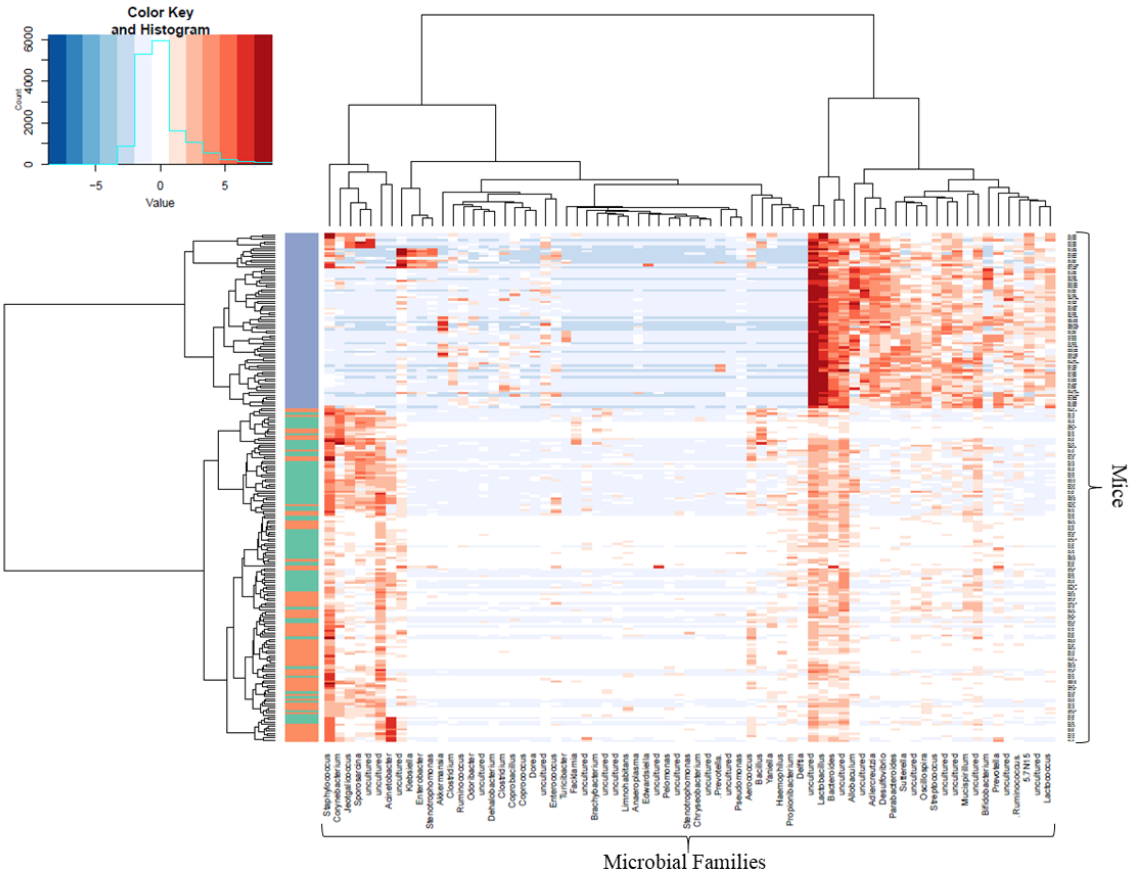

a.
Shannon Diversity
2.82
2.11
0.41
BEM BOM DAVIS FIV GIG GIT HAX2 HAZ JUNIOR LAM LAT LEL LEM LOT LUF LUG NUK POH PUB S129S SAT SEH STUDRY TAS TOFU TOP WAB2 XAV YID ZIF2
Mouse Strains
b.
Color Key and Histogram
Count
Value
Mice
Erysipelotrichaceae
Ruminococcaceae
Enterobacteriaceae
Comamonadaceae
Propionibacteriaceae
Moraxellaceae
Bacteroidaceae
Lachnospiraceae
Lactobacillaceae
S24.7
Staphylococcaceae
Corynebacteriaceae
Aerococcaceae
Planococcaceae
Microbial Families
c.
PC2 (15.5%)
PC1 (34.9%)
Plannococaceae
Aerococcaceae
Lactobacillaceae
Corynebacteriaceae
Moraxellaceae
JUNIOR
LUG
LOT

**Figure 1. An overview of the cutaneous microbiota across 30 different strains of mice from the Collaborative Cross.** (**a**) Scatterplot of alpha diversities for each sample, grouped by mouse strain. Most mice show similar alpha-diversities (Shannon, IQR=0.59, median=2.11) with many strains showing low intra-strain variability. Some strains such as TOFU and LUG show lower diversity scores in all mice within their strain, additionally some strains show much larger intra-strain variabilities, such as TAS and TOP. Diversity of the skin microbiome shows strong strain dependent differences (Kruskal-wallis, *p*-value<0.0001). (**b**) Hierarchical clustering heat-map of `core` microbiota, centred-log ratio read counts of 16S rRNA gene sequencing of healthy dorsal skin swabs. *Staphylococcaceae* was the most abundant bacterial family in the majority of mice, with *Corynebacteriaceae* showing a particularly partisan abundance across different mice. Many mice cluster strongly within their respective strain, with PERMANOVA analysis indicating a significant strain effect (R-squared>0.5, *p*-value=0.001). We cut the hierarchical dendrogram into 6 groups (colours randomly assigned to each group) based on visual inspection resulting in distinct characteristics such as high *Corynebacteriaceae* or *Moraxellaceae* abundance. (**c**) Principal component analysis of centred-log ratio read counts further highlights the differences in abundance of microbiota families between mice. Some strains such as LUG and LOT can be seen clustering very closely together indicating a strong strain effect. With that said, the strain JUNIOR shows that not all mouse strains have a strong preference toward certain microbiota compositions.

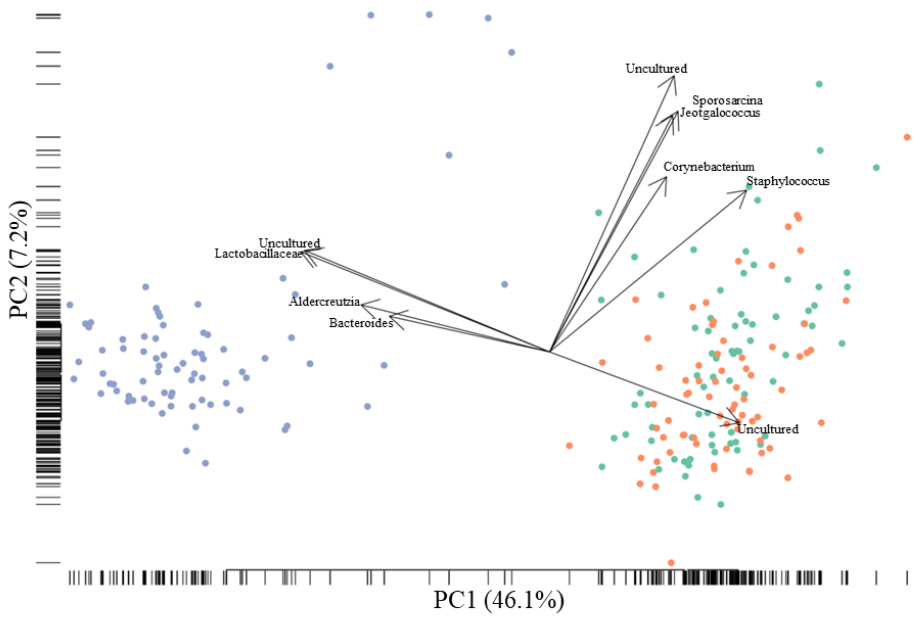
332

**Figure 2.**

**a.**




**b.**




**c.**

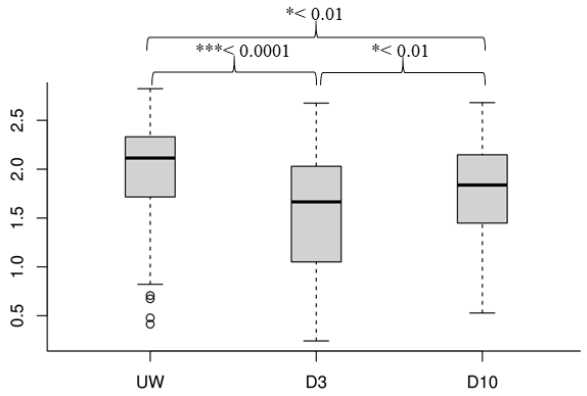
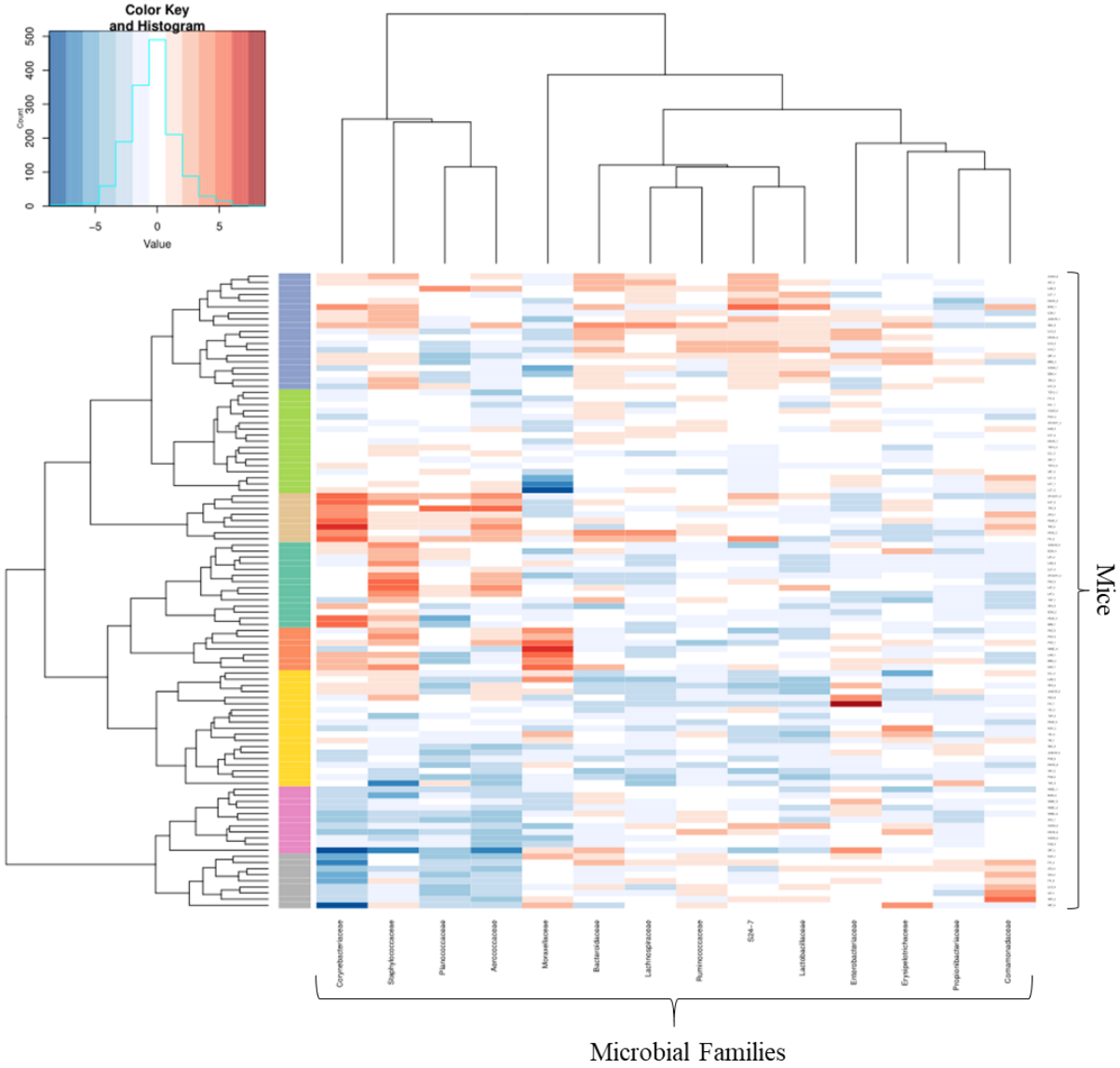


**Figure 2. Comparison of diversity and composition of faecal, unwounded and wounded skin microbiota** (**a**) Hierarchical clustering heatmap of centred-log ratio read counts from all 3 sample microbiota types (faecal = blue, unwounded = green, day-3 post-wounding = orange). Both types of skin microbiota samples show enrichment in certain bacterial genera such as *Staphylococcus*, *Corynebacterium* and *Acinetobacter* compared to faecal samples. (**b**) Principal component analysis of combined sample sets at the genus level. Day-3 post-wound microbiota cluster with unwounded microbiota samples and separately from faecal samples (faecal = blue, unwounded = green, day-3 post-wounding = orange). (**c**) Skin microbial diversity at different time-points (Shannon alpha diversity index). Overall alpha diversity significantly decreases from unwounded to day-3, increasing again by day-10 though remaining significantly depressed compared to unwounded skin.

**Figure 3.**

**a.**


Color Key and Histogram
Count
Value
Mice
Microbial Families

**b.**

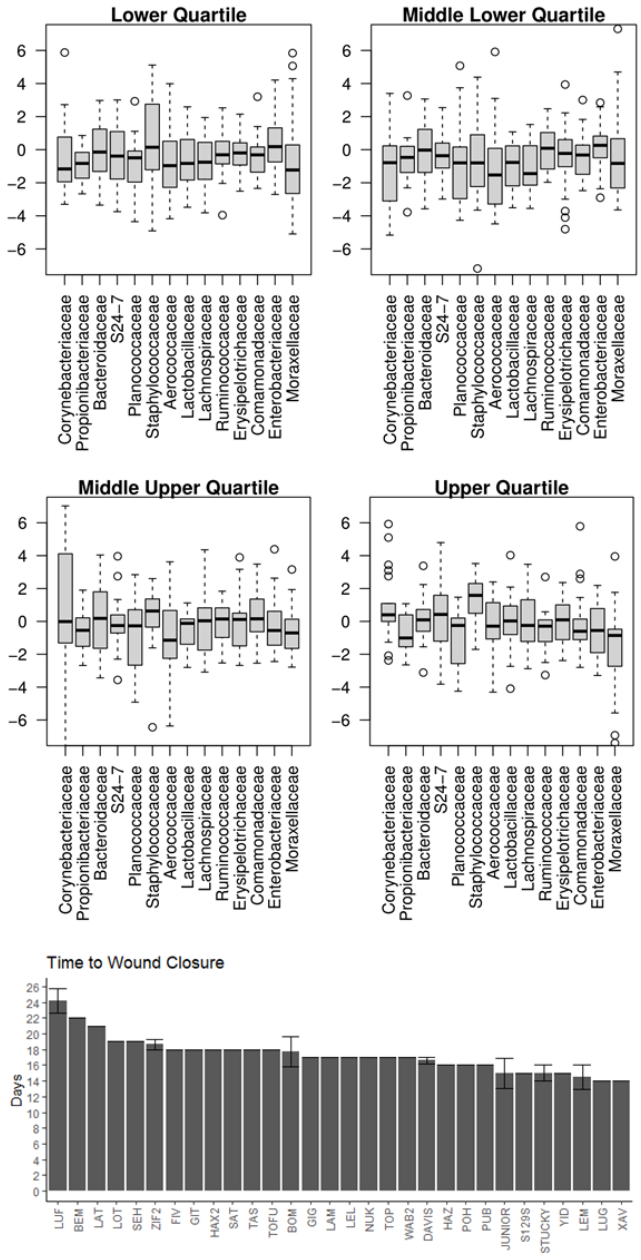
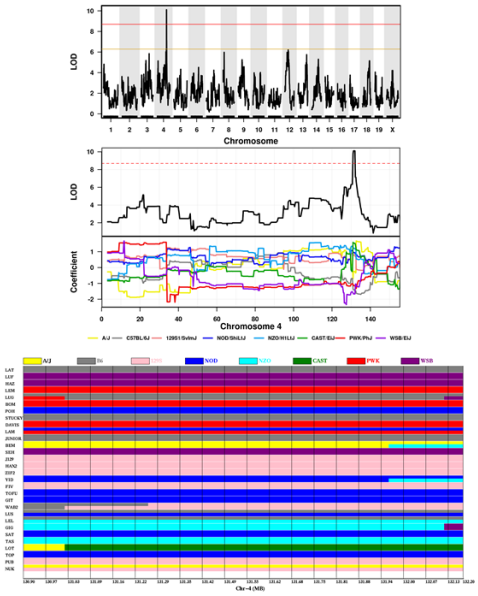

Lower Quartile
Middle Lower Quartile
Middle Upper Quartile
Upper Quartile
Corynebacteriaceae
Propionibacteriaceae
Bacteroidaceae
S24–7
Planococcaceae
Staphylococcaceae
Aerococcaceae
Lactobacillaceae
Lachnospiraceae
Ruminococcaceae
Erysipelotrichaceae
Comamonadaceae
Enterobacteriaceae
Moraxellaceae
Time to Wound Closure
Days
LUF
BEM
LAT
LOT
SEH
ZIF2
FIV
GIT
HAX2
SAT
TAS
TOFU
BOM
GIG
LAM
LEL
NUK
TOP
WAB2
DAVIS
HAZ
POH
PUB
JUNIOR
S129S
STUCKY
YID
LEM
LUG
XAV

**Figure 3. Microbiota composition changes from unwounded skin to day 3 post-wounding.** (**a**) Taking the difference between day 3 post-wounding and unwounded centred-log ratio matrices shows the relative increase/decrease in microbial abundance during the early stages of wound healing. No bacterial families show a consistent pattern of increased/decreased abundance across all mice during wound healing, though many mouse strains show similar within strain patterns of microbiota changes. (**b**) Boxplots of bacterial family centred-log ratio abundance changes during wound healing across all mice, grouped by quartiles of mouse healing speed (Top and Middle). Each bacterial family shows a large spread of values including both mice that increase their relative abundance and mice that decrease. No associations can be seen between a single family of bacteria and faster/slower healing mice. Days to full wound closure across all mouse strains (Bottom).

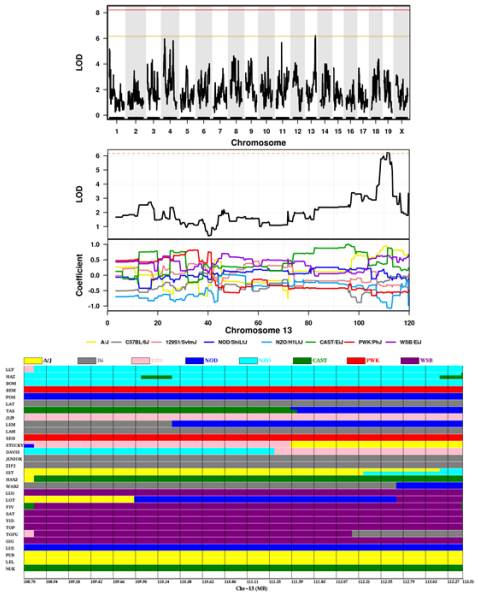
**Supplemental 1.**

361

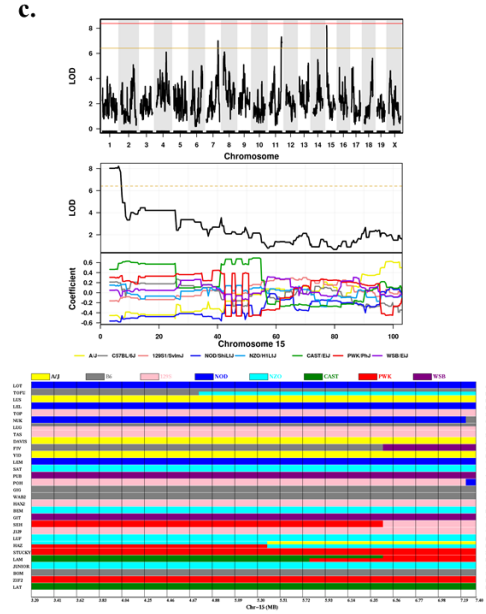
**a.**



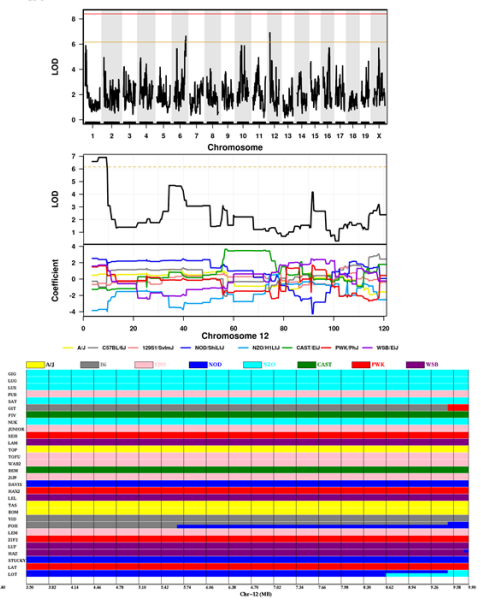

**b.**



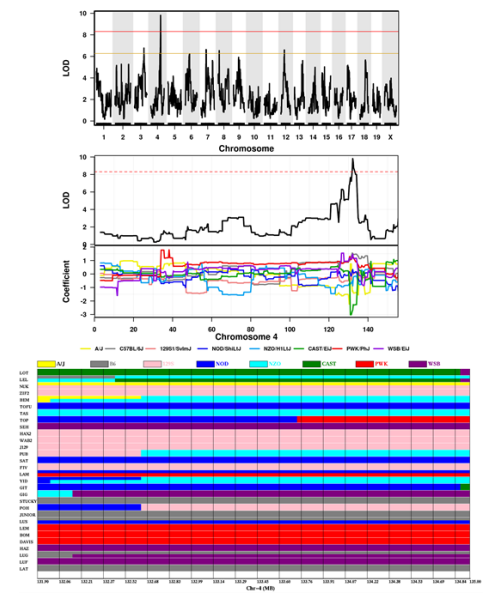

**c.**




**d.**



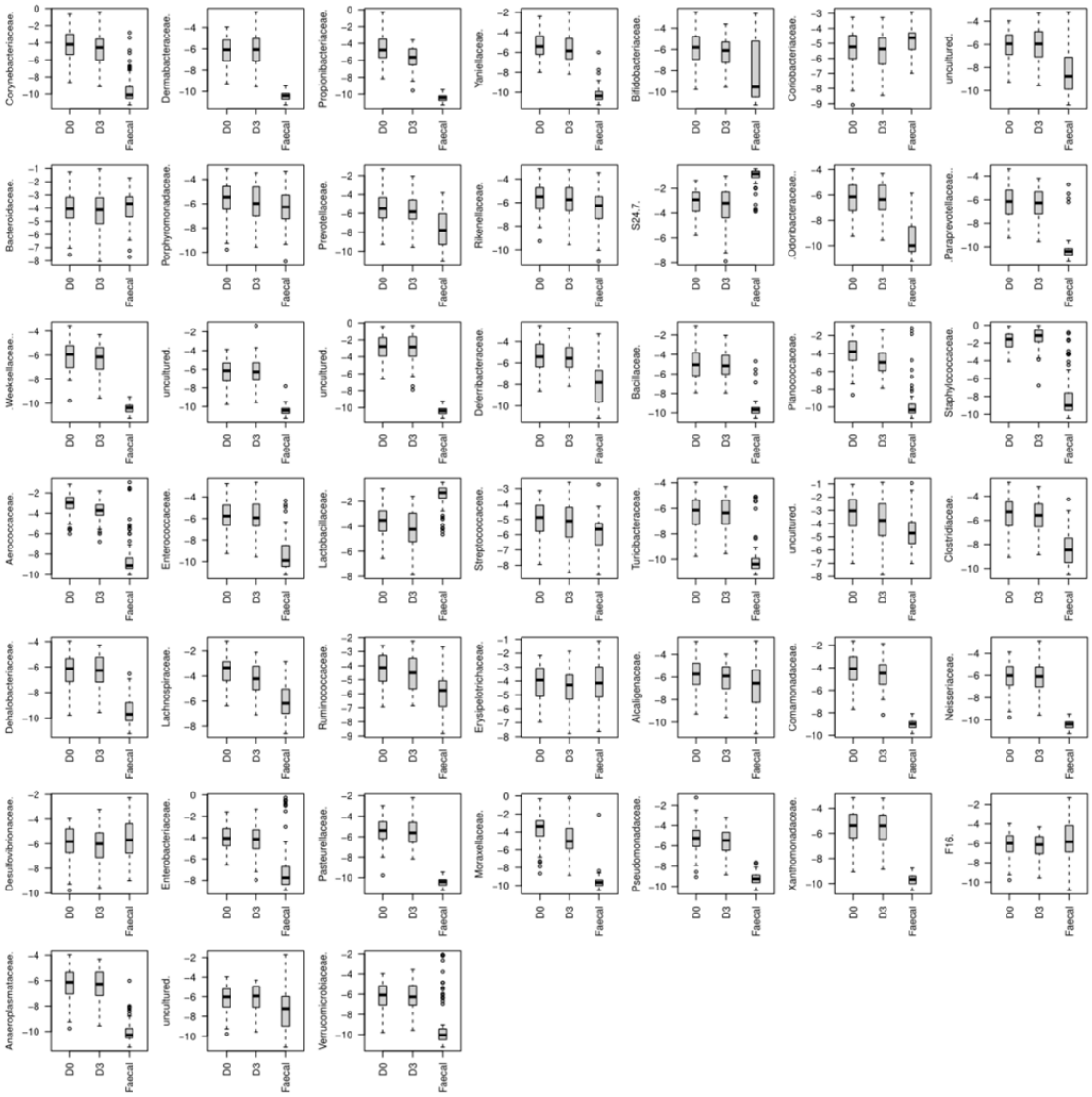

**e.**



**Supplemental 1. GWAS based on haplotype reconstructions of Collaborative Cross mouse genomes.** (**a**) LOD-score based analysis of *Staphylococcus* centred log-ratio abundance inputed to GeneMiner. A significant QTL can be identified on chromosome 4, (-2 LOD score drop region, 130.9-132.2Mbps). Haplotype diagrams indicate the founder strains WSB, CAST and/or PWK may be responsible for the peak. (**b**) LOD-score based analysis of *Aerococcus* centred log-ratio abundance inputed to GeneMiner. A suggestive QTL can be identified on chromosome 13 (-2 LOD score drop region, 108.7-113.51Mbps). Haplotype diagrams indicate the founder strains NZO, AJ and/or WSB may be responsible for the peak. (**c**) LOD-score based analysis of Shannon's diversity scores imputed to GeneMiner. A suggestive QTL can be identified on chromosome 15 (-2 LOD drop region, 3.2-7.4 Mbps). Haplotype diagrams indicate CAST, AJ and/or NOD as likely responsible for the peak. (**d,e**) PCA-based GWAS using principal components 1 and 3 (**d, e**). Suggestive QTL can be found on Chr 12 for both principle components, though in different regions (-2 LOD drop region PC1, 3.5-9.9Mbps, PC3, 47.7-60.4Mbps). For principal component 3 (**e**) a significant peak on Chr4 overlapping the same region as the previous significant peak for *Staphylococcus* can be seen primarily due to the CAST and PWK founder strains (-2 LOD drop region, 131.9-135Mbps).

## Supplemental 2.

**a.** Log-proportional abundances of all bacterial families (including non-core) (70 mice)




**b.**

Table 1: Mice Used

| Strain | Count |
|---|---|
| BEM | 4 |
| BOM | 4 |
| DAVIS | 5 |
| FIV | 5 |
| GIG | 3 |
| GIT | 2 |
| HAX2 | 4 |
| HAZ | 5 |
| JUNIOR | 5 |
| LAM | 4 |
| LAT | 3 |
| LEL | 3 |
| LEM | 2 |
| LOT | 5 |
| LUF | 5 |
| LUG | 4 |
| NUK | 4 |
| POH | 2 |
| PUB | 3 |
| S129S | 5 |
| SAT | 4 |
| SEH | 4 |
| STUCKY | 3 |
| TAS | 3 |
| TOFU | 4 |
| TOP | 5 |
| WAB2 | 5 |
| XAV | 2 |
| YID | 3 |
| ZIF2 | 4 |

**c.**

Table 2: Linear Model of Staphylococcus and Aerococcus Diversity Associations

| Bacteria | Estimate | P.Value |
|---|---|---|
| Staphylococcus | -0.0751578 | 0.0682618 |
| Aerococcus | -0.1168848 | 0.0657272 |

**Supplemental 2.** (**a**) Qualitative summary of all family-level (including non-core skin) log-proportional abundances in unwounded, day 3 post-wound, and faecal samples (70 mice). *Propionibacteriaceae* and *Bacteroidales* family S24-7 show markedly different abundances between skin and faecal samples. (**b**) List of all Collaborative Cross strains used for initial experiment, including mice per strain used (114 total mice). (**c**) Based on the results from differential bacterial genera abundance analysis of most diverse and least diverse mice, we modelled Shannon diversity regressed against *Staphylococcus* and *Aerococcus* CLR abundances. Negative values for estimate indicate lower diversity values for higher CLR values for both *Staphylococcus* and *Aerococcus* with an overall trend across all mice. (**d**) Schematics for the processes of 16S rRNA gene amplicon sequence data analysis.

| Haplotype | Genes | Region_(Chr:Mbps) |
|---|---|---|
| PWK_PhJ | Laptm5,Matn1 | 4:130.9-132.2 |
| CAST_EiJ | Epb4.1,Gm10300,Laptm5,Matn1,Mecr,Ptpru,Srsf4,Tmem200b | 4:130.9-132.2 |
| WSB_EiJ | Epb4.1,Mecr | 4:130.9-132.2 |
| PWK_PhJ | Gpatch3,Rps6ka1 | 4:131.9-135.0 |
| CAST_EiJ | Aim1l,Arid1a,Atpif1,Catsper4,Cnksr1,Epb4.1,Extl1,Eya3,Fam46b, ,Gm10300,Gm7534,Gmeb1,Grrp1,Kdf1,Map3k6,Nr0b2,Nudc,Pdik1l, Phactr4, Pigv,Ptafr,Rab42,Rhd,Rpa2,Rsrp1,Sepn1,Smpdl3b, Srsf4,Stx12,Taf12,Tmem200b,Ubxn11,Xkr8,Zdhhc18,Zfp593,Zfp683 | 4:131.9-135.0 |
| WSB_EiJ | Atpif1,Cd164l2,Epb4.1,Gmeb1,Trnp1 | 4:131.9-135.0 |

**Supplemental 3. Table of candidate genes based on GWAS analysis.** List of genes identified in from the QTL on Chromosome 4 present in 2 different GWAS. Genes are taken from Sanger Mouse SNP Viewer (https://www.sanger.ac.uk/sanger/Mouse_SnpViewer/rel-1505) for the LOD 2 drop region around the QTL peak. A total of 38 genes were identified that possessed SNPs specific to any of the haplotypes WSB, PWK or CAST.

## Materials & Methods

### *Mice*

The Collaborative Cross (CC) program used 8 founding strains of mice to produce several hundred recombinant mouse inbred (RI) strains [20] that were inbred over multiple generations to greater than 90% homozygosity [43]. This high rate of homozygosity ensures that most regions of the genome are defined by the genetic contribution from a single founder strain (haplotype), simplifying analysis. However, the substantial heterogeneity in haplotypes between different strains can result in marked variation in any phenotype across the CC. Greatly reduced costs and complexities can be achieved with CC mice, since they are all genotyped and the founder genome sequences are available at the Sanger Institute (https://www.sanger.ac.uk) [20]. Lastly, the large genetic scope of the CC RI strains, founding strains were chosen to maximise genetic richness, provides a powerful resource to investigate murine genetics of complex biological problems.

All mice from the CC (114) were housed in the UQ Centre for Clinical Research Animal Facility. All animal experimentation was conducted in accordance with institutional ethical requirements and approved by the University of Queensland Animal Ethics Committee. Only female mice were used with the number of mice per strain varying due to availability. Two mice from the strains XAV, GIT, LEM and POH were used whereas, all other strains had 3 or more mice (**Supp.2**).

### *Collection*

Mice were anaesthetized with 2% isoflurane and a sterile rayon swab moistened with TE buffer was used to collect a microbiota sample from a 1.5x1.5cm of dorsal skin, followed by full-thickness excisional wounds of the same area and second swab was immediately used to sample this fresh wound (library sizes too small to analyse). This second swab on the excisional wound served as technical control for contamination for every single mouse. At this initial timepoint a fresh stool sample was collected. Additional swabs were taken at days 3 and 10 post-wounding. All samples were stored in 2ml of TE buffer at -80$^{o}$C for later sequencing.

### *Microbiota community profiling and data analyses*

Microbial DNA was extracted from swab samples of dorsal skin using the Maxwell 16 LEV Buccal Swab DNA kit according to manufacturer's recommendations. The resulting DNA samples were then used to produce bar-coded PCR amplicon libraries of the V6-V8

hypervariable region of the 16S rRNA gene using the universal microbial primers with Illumina primer overhang adapters as follows:

- Forward Primer – 926F:

5’ - TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG AAA CTY AAA KGA ATT GRC GG – 3'

- Reverse Primer – 1392R:

5’ - GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GAC GGG CGG TGW GTR C – 3'

Sequencing used the Illumina MiSeq sequencing platform and protocols developed by the UQ-Australian Centre for Ecogenomics (www.ecogenomic.org). A sequence Phred quality threshold of 20 was used and sequences checked for chimeras using USEARCH version 6.1.544 [44]. Mapping and clustering of reads into operational taxonomic units (OTUs) with 97% identity threshold against Greengenes core set database 13.8 [45], was performed using Quantitative insight into Microbial Ecology (QIIME) version 1.9.1 [46] and PyNast [47]. OTUs were then compiled into an OUT table for further analysis.

***Data Analysis***

Shannon measures of alpha diversities were calculated using the ‘Vegan’ package in R [48, 49] with the function ‘diversity’. The percentage abundance for each bacterial family was calculated and those bacterial families failing a threshold of 0.1% abundance in 50% or more mice were removed, leaving a total of 14 ‘core’ bacterial families remaining in the dataset. Centred log-ratios (CLR) were calculated for all mice using the ‘core’ bacterial families as a sub-composition. Comparisons of murine skin microbiotas similarities were then performed using the PERMANOVA and PCA using Aitchison distance (Euclidean distance after CLR [50].

***QTL Analysis***

Significantly differentially abundant genera based on the upper and lower quartiles of diversity were regressed against mouse genotypes in the GeneMiner software (www.sysgen.org/GeneMiner). GWAS made use of haplotype reconstructions, as detailed previously [51].

We also performed principal component analysis on the centred log-ratios of abundances and regressed the principal components against mouse genotypes. This style of PCA based GWAS has been suggested as a way of identifying pleiotropic QTL [24], due to each component representing a multivariate vector consisting of all phenotypes of interest (in this case all 'core' microbiota families).


## Acknowledgements

We would like to thank Geniad Pty Ltd for the generous provision of CC mice. This work was supported by funding from the National Health and Medical Research Council of Australia (APP1082438). KK is supported by a Queensland government, Advance Queensland Research Fellowship.


## Competing Interests

The authors declare no competing interests.

## References


1. Catinean, A., et al., *Microbiota and Immune-Mediated Skin Diseases-An Overview.* Microorganisms, 2019. **7**(9).
2. Schoeler, M. and R. Caesar, *Dietary lipids, gut microbiota and lipid metabolism.* Rev Endocr Metab Disord, 2019. **20**(4): p. 461-472.
3. Forsythe, P., et al., *Mood and gut feelings.* Brain Behav Immun, 2010. **24**(1): p. 9-16.
4. Kobyliak, N., O. Virchenko, and T. Falalyeyeva, *Pathophysiological role of host microbiota in the development of obesity.* Nutr J, 2016. **15**: p. 43.
5. Pessi, T., et al., *Interleukin-10 generation in atopic children following oral Lactobacillus rhamnosus GG.* Clin Exp Allergy, 2000. **30**(12): p. 1804-8.
6. Kong, H.H., et al., *Temporal shifts in the skin microbiome associated with disease flares and treatment in children with atopic dermatitis.* Genome Res, 2012. **22**(5): p. 850-9.
7. Min, K.R., et al., *Association between baseline abundance of Peptoniphilus, a Gram-positive anaerobic coccus, and wound healing outcomes of DFUs.* PLoS One, 2020. **15**(1): p. e0227006.
8. Wolcott, R.D., et al., *Evaluation of the bacterial diversity among and within individual venous leg ulcers using bacterial tag-encoded FLX and titanium amplicon pyrosequencing and metagenomic approaches.* BMC Microbiol, 2009. **9**: p. 226.
9. Grice, E.A., et al., *Topographical and temporal diversity of the human skin microbiome.* Science, 2009. **324**(5931): p. 1190-2.
10. Grice, E.A. and J.A. Segre, *The skin microbiome.* Nat Rev Microbiol, 2011. **9**(4): p. 244-53.
11. McKnite, A.M., et al., *Murine gut microbiota is defined by host genetics and modulates variation of metabolic traits.* PLoS One, 2012. **7**(6): p. e39191.

12. Benson, A.K., et al., *Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors.* Proc Natl Acad Sci U S A, 2010. **107**(44): p. 18933-8.
13. Scharschmidt, T.C., et al., *Matriptase-deficient mice exhibit ichthyotic skin with a selective shift in skin microbiota.* J Invest Dermatol, 2009. **129**(10): p. 2435-42.
14. Byrd, A.L., et al., *Staphylococcus aureus and Staphylococcus epidermidis strain diversity underlying pediatric atopic dermatitis.* Sci Transl Med, 2017. **9**(397).
15. Grice, E.A., et al., *A diversity profile of the human skin microbiota.* Genome Res, 2008. **18**(7): p. 1043-50.
16. Oh, J., et al., *Temporal Stability of the Human Skin Microbiome.* Cell, 2016. **165**(4): p. 854-66.
17. Gardiner, M., et al., *A longitudinal study of the diabetic skin and wound microbiome.* PeerJ, 2017. **5**: p. e3543.
18. Redel, H., et al., *Quantitation and composition of cutaneous microbiota in diabetic and nondiabetic men.* J Infect Dis, 2013. **207**(7): p. 1105-14.
19. Loesche, M., et al., *Temporal Stability in Chronic Wound Microbiota Is Associated With Poor Healing.* J Invest Dermatol, 2017. **137**(1): p. 237-244.
20. Churchill, G.A., et al., *The Collaborative Cross, a community resource for the genetic analysis of complex traits.* Nat Genet, 2004. **36**(11): p. 1133-7.
21. Ferguson, B., et al., *Melanoma susceptibility as a complex trait: genetic variation controls all stages of tumor progression.* Oncogene, 2015. **34**(22): p. 2879-86.
22. Murtagh, F., Legendre, P., *Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion?* Journal of Classification, 2014. **31**: p. 274-295.
23. Ward, J.H., *Hierarchical Grouping to Optimize an Objective Function.* Journal of the American Statistical Association, 1963. **58**: p. 263-244.
24. Zhang, W., et al., *PCA-Based Multiple-Trait GWAS Analysis: A Powerful Model for Exploring Pleiotropy.* Animals (Basel), 2018. **8**(12).
25. Shukla, S.D., et al., *Platelet activating factor receptor: gateway for bacterial chronic airway infection in chronic obstructive pulmonary disease and potential therapeutic target.* Expert Rev Respir Med, 2015. **9**(4): p. 473-85.
26. Iovino, F., et al., *Signalling or binding: the role of the platelet-activating factor receptor in invasive pneumococcal disease.* Cell Microbiol, 2013. **15**(6): p. 870-81.
27. Stafforini, D.M., et al., *Platelet-activating factor, a pleiotrophic mediator of physiological and pathological processes.* Crit Rev Clin Lab Sci, 2003. **40**(6): p. 643-72.
28. Sahu, R.P., et al., *Topical application of a platelet activating factor receptor agonist suppresses phorbol ester-induced acute and chronic inflammation and has cancer chemopreventive activity in mouse skin.* PLoS One, 2014. **9**(11): p. e111608.
29. Heinz, L.X., et al., *The Lipid-Modifying Enzyme SMPDL3B Negatively Regulates Innate Immunity.* Cell Rep, 2015. **11**(12): p. 1919-28.
30. Velarde, M.C., et al., *Pleiotropic age-dependent effects of mitochondrial dysfunction on epidermal stem cells.* Proc Natl Acad Sci U S A, 2015. **112**(33): p. 10407-12.
31. Keyes, B.E., et al., *Impaired Epidermal to Dendritic T Cell Signaling Slows Wound Repair in Aged Skin.* Cell, 2016. **167**(5): p. 1323-1338 e14.
32. Shettigar, K. and T.S. Murali, *Virulence factors and clonal diversity of Staphylococcus aureus in colonization and wound infection with emphasis on diabetic foot infection.* Eur J Clin Microbiol Infect Dis, 2020. **39**(12): p. 2235-2246.
33. Kong, H.H., *Skin microbiome: genomics-based insights into the diversity and role of skin microbes.* Trends Mol Med, 2011. **17**(6): p. 320-8.

34. Yung, D.B.Y., K.J. Sircombe, and D. Pletzer, *Friends or enemies? The complicated relationship between Pseudomonas aeruginosa and Staphylococcus aureus.* Mol Microbiol, 2021.
35. Khan, A.A., et al., *Polymorphic Immune Mechanisms Regulate Commensal Repertoire.* Cell Rep, 2019. **29**(3): p. 541-550 e4.
36. McKenzie, V.J., et al., *Co-habiting amphibian species harbor unique skin bacterial communities in wild populations.* ISME J, 2012. **6**(3): p. 588-96.
37. Belheouane, M., et al., *Assessing similarities and disparities in the skin microbiota between wild and laboratory populations of house mice.* ISME J, 2020. **14**(10): p. 2367-2380.
38. Ross, A.A., et al., *Comprehensive skin microbiome analysis reveals the uniqueness of human skin and evidence for phylosymbiosis within the class Mammalia.* Proc Natl Acad Sci U S A, 2018. **115**(25): p. E5786-E5795.
39. Roth, S.A., et al., *The pattern recognition receptor NOD2 mediates Staphylococcus aureus-induced IL-17C expression in keratinocytes.* J Invest Dermatol, 2014. **134**(2): p. 374-380.
40. Naik, S., et al., *Commensal-dendritic-cell interaction specifies a unique protective skin immune signature.* Nature, 2015. **520**(7545): p. 104-8.
41. Srinivas, G., et al., *Genome-wide mapping of gene-microbiota interactions in susceptibility to autoimmune skin blistering.* Nat Commun, 2013. **4**: p. 2462.
42. Verbanic, S., et al., *Microbial predictors of healing and short-term effect of debridement on the microbiome of chronic wounds.* NPJ Biofilms Microbiomes, 2020. **6**(1): p. 21.
43. Morgan, A.P. and C.E. Welsh, *Informatics resources for the Collaborative Cross and related mouse populations.* Mamm Genome, 2015. **26**(9-10): p. 521-39.
44. Edgar, R.C., *Search and clustering orders of magnitude faster than BLAST.* Bioinformatics, 2010. **26**(19): p. 2460-1.
45. DeSantis, T.Z., et al., *Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB.* Appl Environ Microbiol, 2006. **72**(7): p. 5069-72.
46. Caporaso, J.G., et al., *QIIME allows analysis of high-throughput community sequencing data.* Nat Methods, 2010. **7**(5): p. 335-6.
47. Caporaso, J.G., et al., *PyNAST: a flexible tool for aligning sequences to a template alignment.* Bioinformatics, 2010. **26**(2): p. 266-7.
48. Jari Oksanen, F.G.B., Michael Friendly, Roeland Kindt, Pierre Legendre, Dan McGlinn, Peter R. Minchin, R. B. O'hara, Gavin L. Simpson, Peter Solymos, M. Henry Stevens, Eduard Szoecs, Helene Wagner, *vegan: Community Ecology Package.* 2019.
49. Team, R.C., *R: A language and environment for statistical computing.* 2020.
50. Gloor, G.B., et al., *Microbiome Datasets Are Compositional: And This Is Not Optional.* Front Microbiol, 2017. **8**: p. 2224.
51. Ram, R. and G. Morahan, *Complex Trait Analyses of the Collaborative Cross: Tools and Databases.* Methods Mol Biol, 2017. **1488**: p. 121-129.