Examining diabetic heel ulcers through an ecological lens: microbial community dynamics associated with healing and infection

Tim J. Sloan[1,2,3*], James C. Turton[1,2], Jess Tyson[1], Alison Musgrove[4], Vicki M. Fleming[2], Michelle M. Lister[2], Matthew W. Loose[1], R. Elizabeth Sockett[1], Mathew Diggle[2], Frances L. Game[4,5], William Jeffcoate[4]

[1] School of Life Sciences, University of Nottingham, Nottingham, UK.

[2] Department of Clinical Microbiology, Queen's Medical Centre, Nottingham University Hospitals NHS Trust, Nottingham, UK.

[3] Path Links Pathology, Northern Lincolnshire and Goole NHS Foundation Trust, Lincolnshire, UK.

[4] Foot Ulcer Trials Unit, Department of Diabetes and Endocrinology, Nottingham University Hospitals NHS Trust, Nottingham, UK.

[5] Department of Diabetes and Endocrinology, Derby Teaching Hospitals NHS Foundation Trust, Derby, UK.

* Corresponding author

Email: tim.sloan@nottingham.ac.uk

Keywords: diabetic foot ulcers, chronic wounds, microbiome, wound healing, polymicrobial interactions

Repositories: Sequences are available at the European Nucleotide Archive www.ebi.ac.uk/ena under study accession reference PRJEB28661.

Subject category: Microbial Ecology and Health

Word count (main body): 5300

Abstract

Purpose: While some microorganisms, such as *Staphylococcus aureus,* are clearly implicated in causing tissue damage in diabetic foot ulcers (DFUs), our knowledge of the contribution of the entire microbiome to clinical outcomes is limited. We profiled the microbiome of a longitudinal sample series of 28 people with diabetes and DFUs of the heel in an attempt to better characterise the relationship between healing, infection and the microbiome.

Methodology: 237 samples were analysed from 28 DFUs, collected at fortnightly intervals for six months or until healing. Microbiome profiles were generated by 16S rRNA analysis, supplemented by targeted nanopore sequencing.

Results/Key findings: DFUs which failed to heal during the study period (20/28, 71.4%) were more likely to be persistently colonised with a heterogeneous community of microorganisms including anaerobes and *Enterobacteriaceae* (log-likelihood ratio 9.56, p=0.008). During clinically apparent infection, a reduction in the diversity of microorganisms in a DFU was often observed due to expansion of one or two taxa, with recovery in diversity at resolution. Modelling of the predicted species interactions in a single DFU with high diversity indicated that networks of metabolic interactions may exist that contribute to the formation of stable communities.

Conclusion: Longitudinal profiling is an essential tool for improving our understanding of the microbiology of chronic wounds, as community dynamics associated with

clinical events can only be identified by examining changes over multiple time points. The development of complex communities, particularly involving *Enterobacteriaceae* and strict anaerobes, may be contributing to poor outcomes in DFUs and requires further investigation.

Introduction

It is generally acknowledged that diabetic foot ulcers (DFUs) present a considerable clinical and economic burden (1, 2). They have multiple and various overlapping causes, including peripheral artery disease (PAD) and different modalities of neuropathy. As intact skin provides a natural barrier, ulceration and exposure of nutrient-rich tissues to the surface predisposes to colonisation with a wide array of microbes. Clinically apparent infection with known pathogens, such as *Staphylococcus aureus*, is recognised as an important cause of deterioration of pre-existing DFUs and delayed healing (3, 4) and best practice is to sample DFUs with clinical signs of infection to identify them (5). However, little is known of interactions which may occur within a diverse community of micro-organisms (and/or with the host) which could result in delayed healing, subclinical tissue damage or predisposition to clinical infection.

Advances in high throughput next generation sequencing (NGS) has enabled the molecular characterisation of entire microbial communities from any particular site, termed the microbiome. Profiling of marker genes, such as the 16S ribosomal RNA gene, or sequencing all the genetic material present in a sample provides an overview of the microbiome - including microorganisms which may be difficult to culture or be present in very low numbers (6). These large datasets can be analysed with data-mining techniques to uncover relationships between the presence of various groups of microorganisms and clinical variables. In a cross-sectional study of soft tissue samples from 40 people with DFUs, Dowd (7) described a broader spread of taxa than previously identified by culture based methods, while in a more recent study of 52

surface swabs from non-infected plantar neuropathic DFUs, Gardner and colleagues (8) reported clustering of identified bacterial taxa into three broad groups, demonstrating significant associations between bacterial groups and DFU depth, area and overall quality of glycaemic control. Further insight was provided by a study of 100 subjects with DFUs from whom repeated samples were obtained in a subpopulation (9). The authors described four different types of bacterial community based on the dominant population of identified bacteria and also studied the extent to which these populations changed. They suggested that more frequent change in community types was a feature of those that healed within 12 weeks of observation.

A deeper understanding of polymicrobial interactions and impact on host tissues could inform risk stratification tools for DFUs and facilitate the design of targeted interventions aimed at altering the colonising microbiota such as phage therapy. In the setting of infection, this may also enable more selective use of antibiotics where clinically appropriate to reduce broad spectrum antibiotic exposure. Sequencing costs, technical limitations and analytical barriers are constraining immediate application of microbiome profiling to the management of DFUs. However, developments in sequencing technologies, particularly those with potential point of care applications such as Oxford nanopore sequencing, promise to minimise these limitations in the foreseeable future (10).

The data presented here were collected as a sub-study of a clinical trial investigating the effectiveness of a simple off-loading device in patients with DFUs affecting the heel (11). Swab samples were taken at the time of recruitment and at fortnightly

intervals up to a maximum of 24 weeks until DFU healing or withdrawal from the parent study. Associations were sought between clinical measures and the microbiota, both at baseline and over the course of the study. Data available from the parent study provided baseline details of the participants as well as clinical information relating to DFU status at each study visit.

Methods

Participants in the parent trial were people with diabetes complicated by DFUs on the heel and were randomised to management either with standard good care plus lightweight fiberglass heel casts or with standard good care alone, and were reviewed in a specialist clinical research service every two weeks. The results of this parent study have been reported in full elsewhere (11). The present study was approved by Derby Research Ethics Committee (IRAS ID 137934). Participating centres were selected from those that were recruiting well to the parent study and which agreed to undertake the additional sampling. All participating patients provided additional informed consent. Since no difference in any outcome measure was observed between the intervention and the usual care arms of the parent study, samples from patients in both groups were combined for the present analyses.

Details of sampling and sample handling

Samples were obtained using flock swabs and the Z technique after cleansing and sharp debridement of the wound surface. Samples were identified only by study centre and study number. Each swab tip was retained in the collection tube and submerged in a PBS-based buffer solution (MoBio Powersoil Collection Fluid, CARLSBAD, CA) to maintain constant pH. For transport, the samples were packed in a protective plastic casing with a cold pack, boxed and sealed in a prepaid envelope and posted first class within 24 hours to a central microbiology laboratory at Nottingham University Hospitals NHS Trust. Samples were subsequently stored at -80 °C until extraction.

DNA extraction

DNA was extracted from the whole sample using the MoBio Powersoil Kit (since renamed to DNeasy PowerSoil Kit) as per manufacturer's guidelines with a pre extraction homogenisation step (4500 rpm, 45 sec x 3). DNA was stored at -80 °C to await subsequent analysis. Obtained DNA was quantified using the Quant-iT™ PicoGreen™ dsDNA Assay Kit (Invitrogen, Waltham, MA) and analysed, in duplicate, on the ABI7500 (Applied Biosystems, Ca, USA).

16S rRNA gene amplification, library preparation and sequencing

Amplicons were generated for the V4 region of the 16S ribosomal RNA gene using the above extracted DNA as a template. Previously validated primers were used to amplify a 359bp amplicon of the V4 hypervariable region of the prokaryotic 16S Ribosomal RNA gene (515F:TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTGCCAGCMGCCGCGGTAA, 806R:GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGGACTACHVGGGTWTCTAAT) (12). 237 out of 256 samples received were successfully amplified.

PCR clean-up was performed using the AMPure XP beads (Beckman Coulter, IN, USA) before proceeding to index PCR, which attached barcodes to individual samples and Illumina sequencing adapters using the Nextera XT Index Kit, as per Illumina (CA, USA) guidelines. Negative controls were included to account for reagent contamination (13). After index PCR, clean-up was repeated as above and amplicons were individually quantified using the Agilent Bioanalyser (CA, USA), prior to normalisation and pooling. The barcoded amplicon pool was run on the Illumina MiSeq

using 2x250bp chemistry. Raw .fastq files were output for subsequent bioinformatics analysis. The sequencing data are publicly available in the European Nucleotide Archive (ENA, acc.no. PRJEB28661).

16S rRNA sequence pre-processing and taxonomic assignment

DNA sequence data was pre-processed with the mare package (14) in R (Foundation for Statistical Computing, Vienna, Austria) as follows: Paired-end reads were merged with quality filtering and trimmed to a uniform length (290 bp). After chimera filtering and clustering with USEARCH (15), reads were assigned to taxonomic labels by alignment with the RDP database (16) using UTAX and a confidence level of 0.6. Taxa below this cut off were annotated as unclassified at that taxonomic level and assigned to the next highest available level. Unique reads occurring at a frequency of less than 1:100,000 were excluded to avoid OTU inflation. Two halophilic genera *Halomonas* and *Shewanella* were identified from negative controls as possible contaminants and excluded from the analysis.

Oxford nanopore library preparation and sequencing

Swab DNA extracts were treated with a NEBNext Microbiome DNA Enrichment Kit (New England Biolabs) to enrich microbial DNA prior to sequencing (17). Library preparation was performed with a Rapid Low Input by PCR Barcoding Kit (SQK-RLB001, Oxford Nanopore Technologies) and libraries sequenced on the MinION platform (Oxford Nanopore Technologies). Base calling was performed with Albacore Sequencing Pipeline Software (version 2.1.2, Oxford Nanopore Technologies) with raw .fastq files output for subsequent analysis.

Direct sample sequencing analysis

MinION reads were screened against the human genome reference hg19 (GenBank assembly GCA_000001405.1) using Minimap2 (18) to remove host DNA contamination. The resulting non-human sequences were annotated with Centrifuge 1.0.3 (19) using the default database for bacteria and archaea. A minimum hit length of 50 and hit length covering at least 5 % of the read were used to filter the resulting annotations. For detection of antimicrobial resistance genes, all of the non-human reads for the sample were assembled using Canu 1.7 (20). The resulting contiguous sequences were assigned a taxonomic identification by Centrifuge and then input to ResFinder 3.0 (21) searching for acquired antimicrobial genes with a 95 % identity threshold.

Statistical and network analysis

Alpha diversity (Inverse Simpson Index, variation within any one sample) and beta diversity (Bray-Curtis Dissimilarity (BCD), variation between samples) were calculated using the vegan package (22) in R (Foundation for Statistical Computing, Vienna), with Principal Coordinates Analysis of BCD and PERMANOVA used for analysis of multivariate microbiota count data. Associations between diversity and clinical categories were tested for using unpaired t-test or ANOVA after assessing for normality with the Shapiro-Wilk method.

To enable analysis of changes occurring within individual DFUs over time, samples were assigned to clusters based on taxonomic profile. Sample clusters were created with the default k-means clustering function in R using relative abundances of taxa at

family level with 150 maximum iterations and 10 simulated starts. After examining the relationship between variance explained and cluster number, six clusters were chosen to best represent the data (Supplementary Fig. 1).

Once samples had been assigned to one of the six clusters, a Markov Model was constructed by counting all cluster transitions over consecutive visits and the counts converted to proportional probabilities by the dividing of total transitions from that cluster. If swab data were unavailable for the subsequent visit, the next available visit was used as the destination cluster. The final cluster for each subject was then used to calculate the probability of transition from each cluster to a healed or unhealed state at study end. Differences in transition probabilities between healed and unhealed DFUs were calculated using a log likelihood ratio test.

A metabolic complementarity network was created from the MinION data by using the NetSeed pairwise metabolic complementarity index (MCI) previously calculated by Levy and Borenstein for 154 species found in the human microbiome (23, 24). Briefly, NetSeed requires as input a metabolic network for each species, based on a list of KEGG reactions per genome available from the Integrated Microbial Genomes project (IMG, http://img.jgi.doe.gov) and generated as described previously (25). The seed set of the network identified by NetSeed represents exogenously acquired compounds which appear as reaction substrates but not products. The MCI is a pairwise comparison of the proportion of the seed set from one species which are products but not seeds in a partner species network, thereby representing complementary by-products. The most closely related species for which an MCI score was available was identified for each of the most abundant taxa in the DFU. These scores were used to create a network with each species represented by a node, and each pairwise MCI

score with a value greater than 0.5 represented by a directed edge towards the species from which benefit is predicted to be derived.

Results

DFU characteristics and microbiota composition

Two hundred and thirty-seven samples were analysed from the index DFU from 28 individuals, of which eight healed during the study period (Supplementary Table 1). The median (IQR) age of the participants was 70.5 (60-77) years. Twenty (71%) were male and 24 (86%) had Type 2 diabetes. The median (IQR) DFU area at randomisation was 161.5 (69-593) mm^2.

Characterisation of the microbiota of each sample by 16S rRNA sequencing identified 63 distinct genera from 37 families, of which the 10 most abundant families accounted for 89.6 % of the overall composition (Supplementary Table 2). The most abundant taxonomic groups were *Corynebacterium* (22.7 %), *Staphylococcus* (15.2 %), unclassified *Enterobacteriaceae* (10.6 %), *Anaerococcus* (6.4 %), *Pseudomonas* (6.1 %) and *Streptococcus* (4.6 %). The majority of the variation between the samples could be explained by subject, indicating that the microbiome of any one DFU tended to be distinct from any other (PERMANOVA, 63.7 % variance explained, $p < 0.001$). There was no correlation between baseline DFU area or NPUAP depth score and microbial diversity (alpha diversity, Inverse Simpson Index, ISI).

Clustering analysis indicated that samples could be separated into six distinct groups based on broad compositional differences, with one or two highly abundant taxa characterising each cluster (Supplementary Fig. 2). Clusters were assigned a letter based on dominant taxa; cluster 1 (n = 20, *Pseudomonaceae* 58.1 % – P), cluster 3 (n = 49, *Corynebacteriaceae* 68.3 % – C), cluster 4 (n = 17, *Micrococcaceae* 87.2 % – M) and cluster 6 (n = 34, *Staphylococcaceae* 74.6 % - S) were relatively homogenous while cluster 2 (n = 61, *Enterobacteriaceae* 33.4 % – E) and cluster 5 (n

= 56, *Clostridiales Incertae Sedis XI* 38.2 % and other anaerobes ~ 20 % - A) exhibited greater taxonomic diversity.

Variations in microbial community dynamics between healing and non-healing DFUs

We sought to investigate whether the microbiota of non-healing DFUs were different to healing DFUs in both overall composition and changes over time.

Of all the 51 samples taken from the 8 DFUs which healed, the majority of samples fell into either the C (17/51, 33.3 %) or S (11/51, 21.6 %) clusters while of the 186 samples taken from 28 unhealed DFUs, the E (52/186, 28.0 %) and A (48/186, 25.8 %) clusters were most frequently represented. Mean baseline alpha diversity in healing DFUs was 2.3 (+/- 0.9) compared to 2.9 (+/- 1.5) for those that did not heal, but this difference was not statistically significant ($p = 0.19$).

For each DFU, the transition in the microbiota between clusters over subsequent visits was used to develop a Markov model (Supplementary Fig. 3 and Supplementary Table 3). The microbiota of individual DFUs tended to remain within the same cluster over subsequent visits with probabilities ranging from 53 % of samples in the S cluster to 71 % in the M cluster. No transitions were observed between the S and P cluster while transitions frequently occurred from the A to E cluster (Probability = 20 %) although less frequently from E to A (8 %).

Specific features of healing and non-healing DFUs were identified from the changes in the microbiota profiles over time. Non-healing DFUs with profiles of either E or A were less likely to transition away from these to other clusters than healing DFUs (probably per visit; non-healing 21 % vs healing 59 %, log likelihood ratio 9.56, p =

0.008). In DFUs which healed, high rates of transition were observed from E towards several clusters, particularly S (33 %) and C (22 %) so that in the final visit before healing, 75 % healed DFUs had a microbiota profile of either S (3/8) or C (3/8) with only 1 each in the E and P clusters and none in either M or A.

Individual profiles from DFUs which were consistent with these overall trends are illustrated in Fig. 1. In Fig. 1(a), the DFU was persistently colonised by a high proportion of *Corynebacterium* and *Staphylococcus* with two episodes of infection treated with co-amoxiclav before finally healing. Fig. 1(b) is an example of one of two DFUs where 3 consecutive samples fell into the E or A clusters with high proportions of anaerobes and *Enterobacteriaceae* observed, accompanied by visible slough, before emergence of *Corynebacterium* and *Staphylococcus* immediately prior to healing. In the profile displayed in Fig. 1(c), *Enterobacteriaceae* and anaerobes are highly abundant during several visits before gradual replacement with *Pseudomonas* pre-healing. Fig. 1(d) is an example of a non-healing DFU where various anaerobes and *Enterobacteriaceae* are present throughout.

Changes in the microbiota associated with infection and antibiotic therapy

Forty-six percent (13/28) of subjects had at least one episode of clinical infection, with a total of 20 discrete episodes of infection identified covering 44 visits.

Infection and antibiotic use often co-occurred such that their relative associations with sample diversity could not be readily distinguished. Overall, samples from infected DFUs exhibited lower microbial diversity than those from uninfected DFU (infected 2.2 +/- 0.9 vs uninfected 3.1 +/- 1.5, $p = 1x10^{-6}$), while diversity also appeared to be lower for visits where subjects were receiving antibiotic therapy (antibiotics 2.4 +/- 1.2 vs no

antibiotics 3.1 +/- 1.5, p = 0.001). Analysing the contribution of both variables indicated that infection was likely to be a greater contributor to the fall in diversity (ANOVA, F = 14.0, p < 0.001) than antibiotic exposure (F = 2.1, p = 0.15).

To further investigate the changes in microbial diversity associated with individual episodes of infection, samples from visits occurring either immediately preceding, during and following an episode of clinical infection were aligned and the microbial diversity plotted**.** This demonstrated a reduction in microbial diversity at the onset of infection with a corresponding recovery in diversity beginning from the final visit where infection was noted clinically (Fig. 2).

Across all samples taken from infected DFUs, cluster E (24/44, 54.5 %) and C (11/44, 25.0 %) were the most prevalent overall, and this was the same for the first visit of each episode of infection (cluster E: 10/20, 50.0 %; cluster C: 4/20, 20.0 %). Despite the recognised importance of *Staphylococcus aureus* as a pathogen in this context, only 1 DFU had a high proportion of *Staphylococcus* at infection onset, falling into the S cluster. There were 3 and 2 episodes of infection beginning with profiles in the P and A clusters, respectively.

Further individual profiles demonstrating temporal changes in the microbiota associated with episodes of clinical infection are shown in Fig. 3. In the DFU profile shown in Fig. 3(a), colonisation at baseline with a mixture of *Kocuria* and other *Micrococcaceae* was observed with an initial alpha diversity index of 1.96. At visit 7, this fell to 1.15 following the emergence and expansion of staphylococci within the DFU, associated with an episode of clinical infection. After commencement of anti-staphylococcal antibiotic therapy (flucloxacillin), the profile reverted to the taxa which were originally dominant.

Several episodes of infection occurred in the DFU shown in Fig. 3(b) which had a high baseline diversity index of 4.79, falling to 1.37 at visit 5 due to an expansion of *Streptococci,* with associated clinical infection from visits 3 to 6. Recovery of microbial diversity to 4.6 at visit 7 with emergence of *Enterobacteriaceae* and later *Pseudomonas* were also observed.

In Fig. 3(c), a DFU colonised with a high proportion of anaerobes over successive visits from 4 to 6 with high diversity (visit 6: 7.81) underwent a dramatic reduction in diversity during an episode of clinical infection at visit 10 (2.58) accompanied by an expansion of *Anaerococcus* and *Peptonophilus*. After a course of co-amoxiclav, the anaerobes were no longer abundant at visit 11, having been replaced by *Enterobacteriaceae*.

The profile shown in Fig. 3(d) shows a DFU colonised by *Enterobacteriaceae, Pseudomonas* and *Corynebacterium*. Onset of infection at visit 8 and empirical treatment with doxycycline was followed later by expansion of *Streptococci* from visit 8 to 9. After a second course of ciprofloxacin and metronidazole, diversity was markedly reduced from 3.8 at visit 9 to 1.1 at visit 11 as *Corynebacterium* replaced the other taxa to become the dominant organism in the DFU.

In addition to the cases described above, abrupt changes were also noted in certain DFUs without any obvious cause such as antibiotic exposure or clinical infection. In other cases the clinical diagnosis of infection was documented but there were no obvious associated changes in the microbiota (Supplementary Fig. 4).

Direct sample sequencing identifies individual species and resistance genes

Short-read sequencing of 16S rRNA amplicons enables high throughput community profiling but is limited by taxonomic resolution and does not provide any additional information about the genetic capability of a community. To complement the 16S profiles, we conducted longer read sequencing using the Oxford Nanopore MinION on the swab DNA extracts for the profile shown in Fig. 3(b). This profile was chosen as it was one of the most diverse profiles containing a mixture of taxa including *Streptococci, Enterobacteriaceae*, *Pseudomonas* and several strict anaerobes.

Seven samples were sequenced from alternate visits (1, 3, 5, 7, 9, 11, and 13). Despite a microbiome enrichment step, only 3 samples yielded > 1000 reads once host DNA sequences had been discarded (V3 16900, V5 2353, V9 2127). After performing sequence annotation with the taxonomic classifier, Centrifuge, (19) and applying strict cut-offs to the output to optimise the accuracy of taxonomic identification, only a small proportion of reads were eventually assigned to a bacterial species (V3 1548, V5 426, V9 172). Taxa identified from V3 compared well to the 16S profile (Fig. 4a). While read numbers were low for V5 and V9, the species identified were also consistent with the 16S results, including *Streptococcus anginosus* (V5, V9)*, Escherichia coli* (V5, V9), *Enterococcus faecalis* (V5)*, Staphylococcus aureus* (V5), *Bacteroides fragilis* (V9) and *Porphyromonas asaccharolytica* (V9).

Analysing the V3 sample for the presence of antimicrobial resistance genes using the ResFinder tool (21) identified *tet*, *erm* and *cfx* genes. Assembling the sequences from this sample and performing taxonomic annotation on the resulting assembly enabled the matching of *tetM, ermA* and *ermB* to *Streptococcus anginosus* and *cxfA* and *tetQ* to *Prevotella intermedia* based on the presence of these genes in continuous DNA sequence with a high quality taxonomic assignment.

Finally, using the more detailed taxonomic profile created from this DFU, we compiled a network of potential inter-species interactions based on a metabolic complementarity index (MCI) previously calculated by Levy *et al*. (23) using the NetSeed algorithm (24) (Fig. 4b). Higher scores on the MCI indicate a species is predicted to derive metabolic benefit from the metabolic by-products of the partner species. The strict anaerobes, particularly *Porphyromonas asaccharolytica,* are predicted to derive benefit from a number of other species, particularly *E. coli* and *P. aeruginosa*, whereas *E. coli* is predicted to confer benefit on all the strict anaerobes, *E. faecalis* and *S. anginosus*, but may not itself derive much metabolic benefit from the presence of these species.

Discussion

We have characterised the variation in the microbiota of 28 DFUs using a high-resolution longitudinal sampling approach over a 6 month period, or until DFU healing. Successive longitudinal sampling is particularly valuable in this context as we were able to observe changes in the DFU microbiota occurring over short time frames and in conjunction with clinical events such as impending healing, infection or antibiotic administration.

The most abundant taxa identified in this study were similar to those found in other molecular studies of DFUs, although the proportions vary between studies (8, 9, 26). We observed a dominance of Gram positive organisms including *Corynebacteria* (22.7 %), *Staphylococci* (15.2 %) and anaerobic members of the families *Clostridiales Incertae Sedis X* (total 13.4 %, including *Anaerococcus*, *Finegoldia* and *Peptinophilus*) and *Micrococcaceae* (7.9 %), with a lower abundance of Gram negative organisms; predominantly the family *Enterobacteriaceae* (11.5 %) and *Pseudomonas* (6.1 %). By comparison, Wolcott *et al.* (26) identified *Staphylococcus aureus* (15.0 %) and *Staphylococcus epidermidis* (10.7 %) as the most abundant species in 910 DFUs of all types using 16S amplicon pyrosequencing with high proportions of *Pseudomonas aeruginosa, Corynebacterium spp.*, *Enterococcus spp.*, *Finegoldia magna* and *Anaerococcus vaginalis* also present. The proportions observed here are likely to reflect the patient group and DFU type selected for the parent study.

There are likely to be a large number of factors which are responsible for shaping the microbiota of an individual DFU. These will include bacteria already present on intact skin and elsewhere in the body prior to ulceration, as well as the microenvironment influenced by oxygen tension, glycaemic control, chronicity of the wound and historical antibiotic exposure (8, 27, 28). This would explain why samples from any one DFU tended to be similar to each other, but often varied greatly in composition to other DFUs. By contrast, the gut microbiota are heavily influenced by dietary intake and tends to be composed of a core group of taxa in individuals who share a similar diet (29).

To overcome the challenge of comparing a heterogeneous group of samples from a limited DFU population, we used a clustering approach similar to that employed in other molecular studies of DFUs (8, 9). Clustering at family level facilitated the grouping of organisms with overlapping genomic content and growth requirements such as the variety of closely related anaerobes within the *Clostridiales Incertae Sedis X*. The resulting six clusters were relatively homogeneous, consistent with the most abundant taxa identified and captured a significant proportion of the overall variation in the data. Repeat sampling from individual DFUs yielded profiles from the same cluster on at least 50 % of visits, consistent with the observation that individual DFUs tended to have unique profiles and retained the same combinations of taxa over time.

Impaired healing of DFUs is influenced by a variety of factors notably ischaemia and neuropathy, but also subtle cellular deficits, particularly impaired leucocyte function and alterations in the normal inflammatory response (4, 30-32). There is a growing

403 body of evidence that the composition of the chronic wound microbiome may also
404 have an impact, but the possible mechanisms for this are not well understood. In a
405 murine wound model comparing wild-type to diabetic mice, Grice *et al.* (33) observed
406 delayed healing in diabetic mice with differences in gene expression associated with
407 the immune response, and increased abundance of *Staphylococcus*, *Aerococcus* and
408 taxa within the *Enterobacteriaceae* and *Porphyromonadaceae*. In a compelling study,
409 Wolcott *et al*. (34) sampled the microbiota from 43 chronic wounds and observed that
410 83% of these interfered with healing when transplanted into a murine chronic wound
411 model.

412

413 In this study, we observed a tendency towards persistent colonisation with a
414 heterogeneous population of bacteria including *Enterobacteriaceae* and several
415 groups of Gram positive and Gram negative anaerobes in DFUs which failed to heal.
416 DFUs which healed tended to be colonised with high proportions of Gram positive
417 aerobes, particularly staphylococci and corynebacteria, prior to healing. An obvious
418 explanation for these findings is that the microbiota are simply determined by the DFU
419 environment; chronic non-healing wounds promoting the growth of organisms
420 preferring moist sites while DFUs which have almost healed supporting organisms
421 commonly found on intact skin. The alternative possibility that must also be considered
422 however, particularly in light of other studies in this area, is that colonisation with
423 certain combinations of microbes may itself be a contributor to delayed healing. In a
424 murine chronic wound model, Dalton *et al.* (35) found that a polymicrobial mixture of
425 Gram positive and Gram negative aerobes and anaerobes was associated with
426 delayed healing compared to wounds colonised with a single organism. Loesche *et al*.
427 (9) also studied longitudinal profiles of the microbiota in DFUs and while they did not

identify specific groups of bacteria associated with poor healing, they also observed that non-healing DFUs tended to have a stable microbiome with less variation over time than healing DFUs.

In addition to environmental factors, competition for nutrients and metabolic synergy between microbes is another key determinant of community composition (36). Analysing the predicted metabolic overlap between species co-occurring in one DFU in this study indicated that strict anaerobes are the likely principal beneficiaries of such metabolic overlap. Anaerobes such as *Porphyromonas* benefit from poor tissue oxygenation in diabetes and possibly also increased levels of metabolic substrates such as alpha-ketoglutarate (37, 38). They may also benefit from micro-anaerobic climates created by other bacteria lowering redox potential, forming surface layers including biofilm and exchange of metabolic by-products. Studies of the microbiome in periodontal disease have already demonstrated the importance of groups of organisms, including anaerobes, working together to impair the host response, degrade connective tissue and promote chronic inflammatory states (39).

Diversity of individual samples tended to be lower during episodes of clinical infection and on antibiotic therapy. In some instances this appeared to indicate the proliferation of potentially pathogenic organisms which may have resulted in the clinical symptoms observed. Alternatively, it could be that inflamed wounds represent a changed local environment in which certain bacteria are able to flourish while others are not. Certain bacteria, such as *E. coli,* are able to use inflammatory mediators, for example nitric oxide, as metabolic substrates potentially giving them a survival advantage in pro-

inflammatory environments (40). Interactions between microbes may result in the promotion or suppression of certain species with virulence potential. Another possible mechanism by which community dynamics may influence the onset of infection is through modulation of virulence potential by inter-species interaction. Ramsey *et al*. observed that *Corynebacterium striatum* was capable of interfering with quorum sensing in *Staphylococcus aureus* with the effect of suppressing virulence factor production (41). While high proportions of staphylococci were only observed in 1 in 20 cases of developing infection, it is possible that proliferation of other species may have altered virulence factor expression by less abundant pathogens such as *Staphylococcus aureus,* resulting in tissue damage, inflammation and clinically apparent infection.

16S rRNA amplicon sequencing enables the identification of organisms that are difficult to isolate with conventional culture, while giving an estimate of the proportion of bacterial taxa present in the sample. Despite potential high throughput and low cost, the principal limitation of short-read sequencing is that taxa cannot be reliably identified to the lowest taxonomic levels, particularly species but also genus in some cases. As higher level taxonomic groupings, such as the family *Enterobacteriaceae*, can contain a wide range of organisms with greatly differing virulence potential, this is a significant drawback. Nonetheless, this form of molecular survey is capable of yielding important insights into the bacterial community dynamics of a DFU and how this relates to clinical events, more so than phenotypic testing. Microbiota profiles were often similar over multiple visits, both in the most abundant taxa identified and their relative proportions. This suggests that the approaches used for sampling, extraction, amplification and sequencing were likely to be relatively consistent between samples.

To complement the 16S analysis, we also analysed swab DNA extracts directly with Oxford nanopore sequencing technology, which is the first time to our knowledge that this has been applied to diabetic wounds. Sequencing the total pool of the DNA in a sample enabled greater taxonomic resolution and identification of antibiotic resistance genes. The amount of information which could be determined from a sample appeared to be directly proportional to the ratio of host to bacterial DNA recovered, as any host DNA sequenced reduced bacterial read depth. The high proportion of host DNA in many of the samples limited a comprehensive characterisation of the microbiota. Technical challenges to be overcome include the optimisation of sample extraction methods and reduction in the proportion of host DNA, however the capacity for rapid sequencing is an appealing prospect for diagnostic utility.

Given the longitudinal study design and need for fortnightly sampling, ulcer base swabs were chosen over tissue sampling. Aside from drawbacks of the sequencing technologies used, this is another potential study limitation as surface swabs may not detect the invasion and proliferation of bacteria in tissue layers during infection. In a cross sectional study of wound swab versus tissue sampling, Nelson *et al*. detected a higher proportion of tissue samples containing at least one recognised pathogen (86.1 %) vs. swabs (70.1 %) with fewer less pathogenic organisms in tissue (42). Host DNA proportions are likely to be even higher in tissue samples than from swabs, however longitudinal microbiota characterisation from tissue samples could potentially give clearer insight into the dynamics of infection if these considerable practical and technical difficulties can be overcome.

This study has yielded important information of the microbiology of difficult to heal DFUs of the heel, including differences in microbial diversity and changes over time in infected, healing and non-healing DFUs. While heel ulcers are a subtype of DFUs, they represent a wide range of DFU types from the small and superficial to the deep and anaerobic. We anticipate that building on this and similar studies by analysing a greater number and employing a wide range of sequencing strategies, including full metagenomics and transcriptional analysis, will add to our understanding of the mechanisms underlying non-healing DFUs with implications for future basic and translational research studies in this field.

Author statements

Author contributions

Conceptualisation, TJS, JCT, MD, VMF, FLG, WJ; Methodology, TJS, JCT, VMF; Investigation, TJS, JCT, VMF, JT, AM, MWL; Visualization, TJS; Writing – Original Draft, TJS, WJ; Writing – Review & Editing – TJS, JCT, VMF, JT, MML, MD, FLG, WJ; Funding Acquisition, RES, MD, FLG, WJ; Resources, JCT, AM, VMF, MML, MWL, RES, MD, FLG, WJ; Supervision, RES, MD, FLG, WJ.

Funding information

The parent trial ISRCTN62524796 ‘Evaluation of lightweight fibreglass heel casts in the management of ulcers of the heel in diabetes’ was funded by the National Institute for Health Research Health Technology Assessment 09/01/53.

This study was funded by research funds held by the Foot Ulcer Trials Unit and Department of Clinical Microbiology at Nottingham University Hospitals NHS Trust.

Acknowledgements

The support of the parent study team, including the Nottingham Clinical Trials Unit, was essential in facilitating this study. In particular, Mara Sprengel and Vivienne Turtle-Savage facilitated the matching of data from the clinical trial with sample data. Illumina and Oxford Nanopore Sequencing was carried out by Deep Seq, School of Life Sciences, University of Nottingham. Discussions with Dr. Bonnie Hurwitz, University of Arizona, were also invaluable in refining the Markov model and other data

visualisations. We also thank the clinical sites which collected and promptly delivered the swab samples.

Conflicts of interest

The authors report no conflicts of interest.

Ethical statements

All participants gave written, informed consent for inclusion in the parent trial, and for molecular analysis to be performed on DFU swabs. Researchers working on sample analysis were given participant data in pseudo-anonymised form by study identifier. Any sequences identified as human from direct sample sequencing were counted and discarded prior to downstream analysis.

References

1. Boulton AJ, Vileikyte L, Ragnarson-Tennvall G, Apelqvist J. The global burden of diabetic foot disease. Lancet. 2005;366(9498):1719-24.
2. Kerr M. Diabetic footcare in England, An economic case study. Diabetes UK. 2017 [Available from: https://www.diabetes.org.uk/Upload/Shared%20practice/Diabetic%20footcare%20in%20England,%20An%20economic%20case%20study%20(January%202017).pdf
3. Prompers L, Schaper N, Apelqvist J, Edmonds M, Jude E, Mauricio D, et al. Prediction of outcome in individuals with diabetic foot ulcers: focus on the differences between individuals with and without peripheral arterial disease. The EURODIALE Study. Diabetologia. 2008;51(5):747-55.
4. Ince P, Abbas ZG, Lutale JK, Basit A, Ali SM, Chohan F, et al. Use of the SINBAD classification system and score in comparing outcome of foot ulcer management on three continents. Diabetes Care. 2008;31(5):964-7.
5. Lipsky BA, Aragon-Sanchez J, Diggle M, Embil J, Kono S, Lavery L, et al. IWGDF guidance on the diagnosis and management of foot infections in persons with diabetes. Diabetes Metab Res Rev. 2016;32 Suppl 1:45-74.
6. Mizrahi-Man O, Davenport ER, Gilad Y. Taxonomic classification of bacterial 16S rRNA genes using short sequencing reads: evaluation of effective study designs. PLoS One. 2013;8(1):e53608.
7. Dowd SE, Wolcott RD, Sun Y, McKeehan T, Smith E, Rhoads D. Polymicrobial nature of chronic diabetic foot ulcer biofilm infections determined using bacterial tag encoded FLX amplicon pyrosequencing (bTEFAP). PLoS One. 2008;3(10):e3326.
8. Gardner SE, Hillis SL, Heilmann K, Segre JA, Grice EA. The neuropathic diabetic foot ulcer microbiome is associated with clinical factors. Diabetes. 2013;62(3):923-30.
9. Loesche M, Gardner SE, Kalan L, Horwinski J, Zheng Q, Hodkinson BP, et al. Temporal Stability in Chronic Wound Microbiota Is Associated With Poor Healing. J Invest Dermatol. 2017;137(1):237-44.
10. Jain M, Olsen HE, Paten B, Akeson M. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biol. 2016;17(1):239.
11. Jeffcoate W, Game F, Turtle-Savage V, Musgrove A, Price P, Tan W, et al. Evaluation of the effectiveness and cost-effectiveness of lightweight fibreglass heel casts in the management of ulcers of the heel in diabetes: a randomised controlled trial. Health Technol Assess. 2017;21(34):1-92.
12. Gilbert JA, Field D, Swift P, Thomas S, Cummings D, Temperton B, et al. The taxonomic and functional diversity of microbes at a temperate coastal site: a 'multi-omic' study of seasonal and diel temporal variation. PLoS One. 2010;5(11):e15545.
13. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. 2014;12:87.
14. Korpela K. mare: Microbiota Analysis in R Easily. R package version 1.0. Zenodo. http://doi.org/10.5281/zenodo.50310 2016.
15. Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat Methods. 2013;10(10):996-8.

16. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, et al. Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res. 2014;42(Database issue):D633-42.
17. Feehery GR, Yigit E, Oyola SO, Langhorst BW, Schmidt VT, Stewart FJ, et al. A method for selectively enriching microbial DNA from contaminating vertebrate host DNA. PLoS One. 2013;8(10):e76096.
18. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018.
19. Kim D, Song L, Breitwieser FP, Salzberg SL. Centrifuge: rapid and sensitive classification of metagenomic sequences. Genome Res. 2016;26(12):1721-9.
20. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 2017;27(5):722-36.
21. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, et al. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother. 2012;67(11):2640-4.
22. Dixon P. VEGAN, a package of R functions for community ecology. Journal of Vegetation Science. 2003;14(6):927-30.
23. Levy R, Borenstein E. Metabolic modeling of species interaction in the human microbiome elucidates community-level assembly rules. Proc Natl Acad Sci U S A. 2013;110(31):12804-9.
24. Carr R, Borenstein E. NetSeed: a network-based reverse-ecology tool for calculating the metabolic interface of an organism with its environment. Bioinformatics. 2012;28(5):734-5.
25. Borenstein E, Kupiec M, Feldman MW, Ruppin E. Large-scale reconstruction and phylogenetic analysis of metabolic environments. Proc Natl Acad Sci U S A. 2008;105(38):14482-7.
26. Wolcott RD, Hanson JD, Rees EJ, Koenig LD, Phillips CD, Wolcott RA, et al. Analysis of the chronic wound microbiota of 2,963 patients by 16S rDNA pyrosequencing. Wound Repair Regen. 2016;24(1):163-74.
27. Smith K, Collier A, Townsend EM, O'Donnell LE, Bal AM, Butcher J, et al. One step closer to understanding the role of bacteria in diabetic foot ulcers: characterising the microbiome of ulcers. BMC Microbiol. 2016;16:54.
28. Gontcharova V, Youn E, Sun Y, Wolcott RD, Dowd SE. A comparison of bacterial composition in diabetic ulcers and contralateral intact skin. Open Microbiol J. 2010;4:8-19.
29. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, et al. Diet rapidly and reproducibly alters the human gut microbiome. Nature. 2014;505(7484):559-63.
30. Blakytny R, Jude E. The molecular biology of chronic wounds and delayed healing in diabetes. Diabet Med. 2006;23(6):594-608.
31. Zykova SN, Jenssen TG, Berdal M, Olsen R, Myklebust R, Seljelid R. Altered cytokine and nitric oxide secretion in vitro by macrophages from diabetic type II-like db/db mice. Diabetes. 2000;49(9):1451-8.
32. Sindrilaru A, Peters T, Wieschalka S, Baican C, Baican A, Peter H, et al. An unrestrained proinflammatory M1 macrophage population induced by iron impairs wound healing in humans and mice. J Clin Invest. 2011;121(3):985-97.
33. Grice EA, Snitkin ES, Yockey LJ, Bermudez DM, Liechty KW, Segre JA, et al. Longitudinal shift in diabetic wound microbiota correlates with prolonged skin defense response. Proc Natl Acad Sci U S A. 2010;107(33):14799-804.

34. Wolcott R, Sanford N, Gabrilska R, Oates JL, Wilkinson JE, Rumbaugh KP. Microbiota is a primary cause of pathogenesis of chronic wounds. J Wound Care. 2016;25(Sup10):S33-S43.
35. Dalton T, Dowd SE, Wolcott RD, Sun Y, Watters C, Griswold JA, et al. An in vivo polymicrobial biofilm wound infection model to study interspecies interactions. PLoS One. 2011;6(11):e27317.
36. Zelezniak A, Andrejev S, Ponomarova O, Mende DR, Bork P, Patil KR. Metabolic dependencies drive species co-occurrence in diverse microbial communities. Proc Natl Acad Sci U S A. 2015;112(20):6449-54.
37. Milner P, Batten JE, Curtis MA. Development of a simple chemically defined medium for Porphyromonas gingivalis: requirement for alpha-ketoglutarate. FEMS Microbiol Lett. 1996;140(2-3):125-30.
38. Tan Q, Wang W, Yang C, Zhang J, Sun K, Luo HC, et al. alpha-ketoglutarate is associated with delayed wound healing in diabetes. Clin Endocrinol (Oxf). 2016;85(1):54-61.
39. Costalonga M, Herzberg MC. The oral microbiome and the immunobiology of periodontal disease and caries. Immunol Lett. 2014;162(2 Pt A):22-38.
40. Winter SE, Lopez CA, Baumler AJ. The dynamics of gut-associated microbial communities during inflammation. EMBO reports. 2013;14(4):319-27.
41. Ramsey MM, Freire MO, Gabrilska RA, Rumbaugh KP, Lemon KP. Staphylococcus aureus Shifts toward Commensalism in Response to Corynebacterium Species. Front Microbiol. 2016;7:1230.
42. Nelson A, Wright-Hughes A, Backhouse MR, Lipsky BA, Nixon J, Bhogal MS, et al. CODIFI (Concordance in Diabetic Foot Ulcer Infection): a cross-sectional study of wound swab versus tissue sampling in infected diabetic foot ulcers in England. BMJ Open. 2018;8(1):e019437.

Abbreviations

ANOVA	Analysis of Variance
bp	Base pairs
BCD	Bray-Curtis Dissimilarity
DFU	Diabetic foot ulcer
IMG	Integrated Microbial Genomes
ISI	Inverse Simpson Index
MCI	Metabolic Complementarity Index
NPUAP	National Pressure Ulcer Advisory Panel
OTU	Operational Taxonomic Unit
PERMANOVA	Permutational Analysis of Variance
PBS	Phosphate buffered saline

Tables and figures

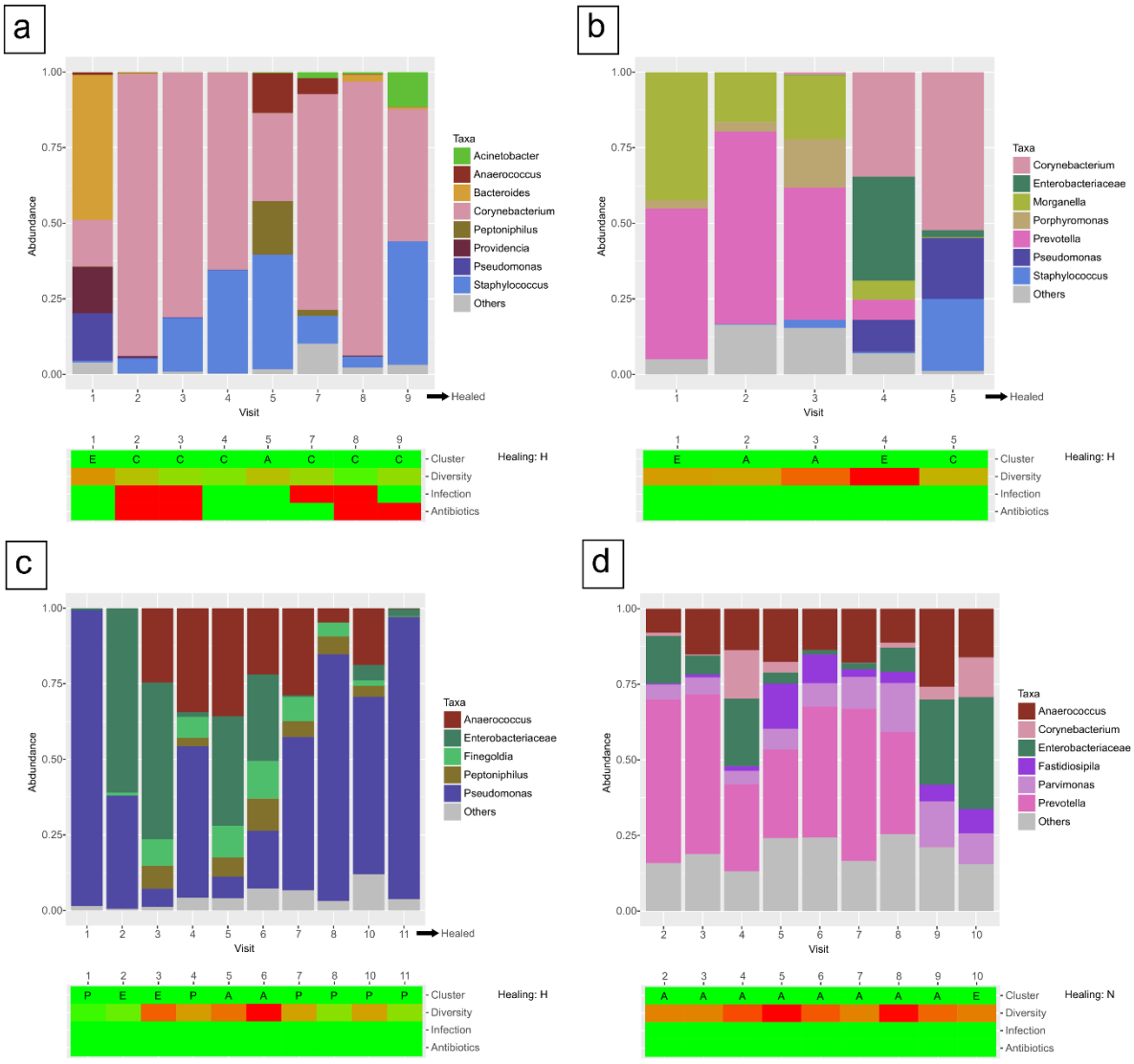
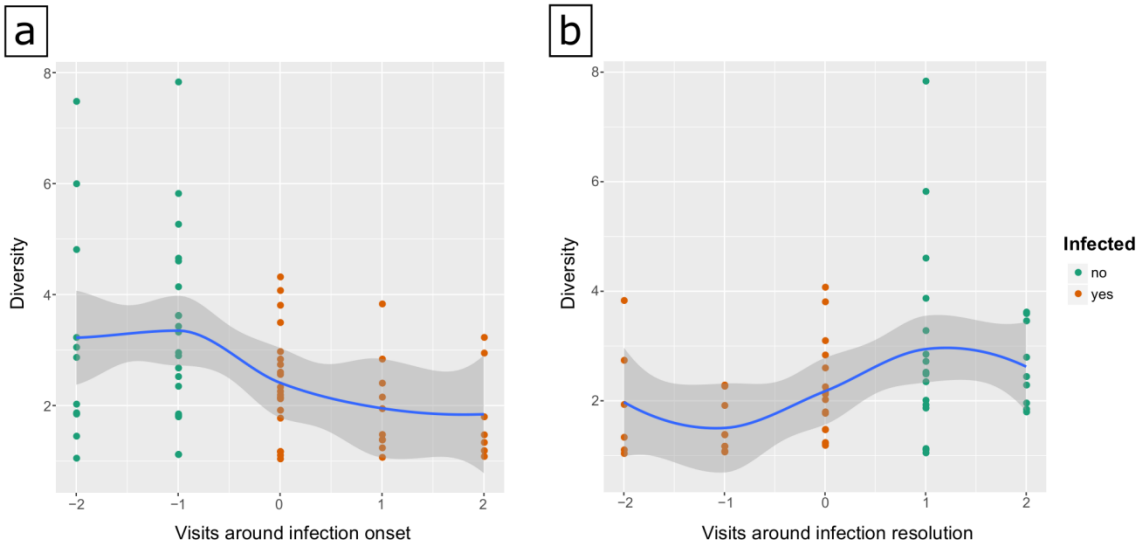


Figure 1. Examples of DFU profiles where healing occurred during the study.

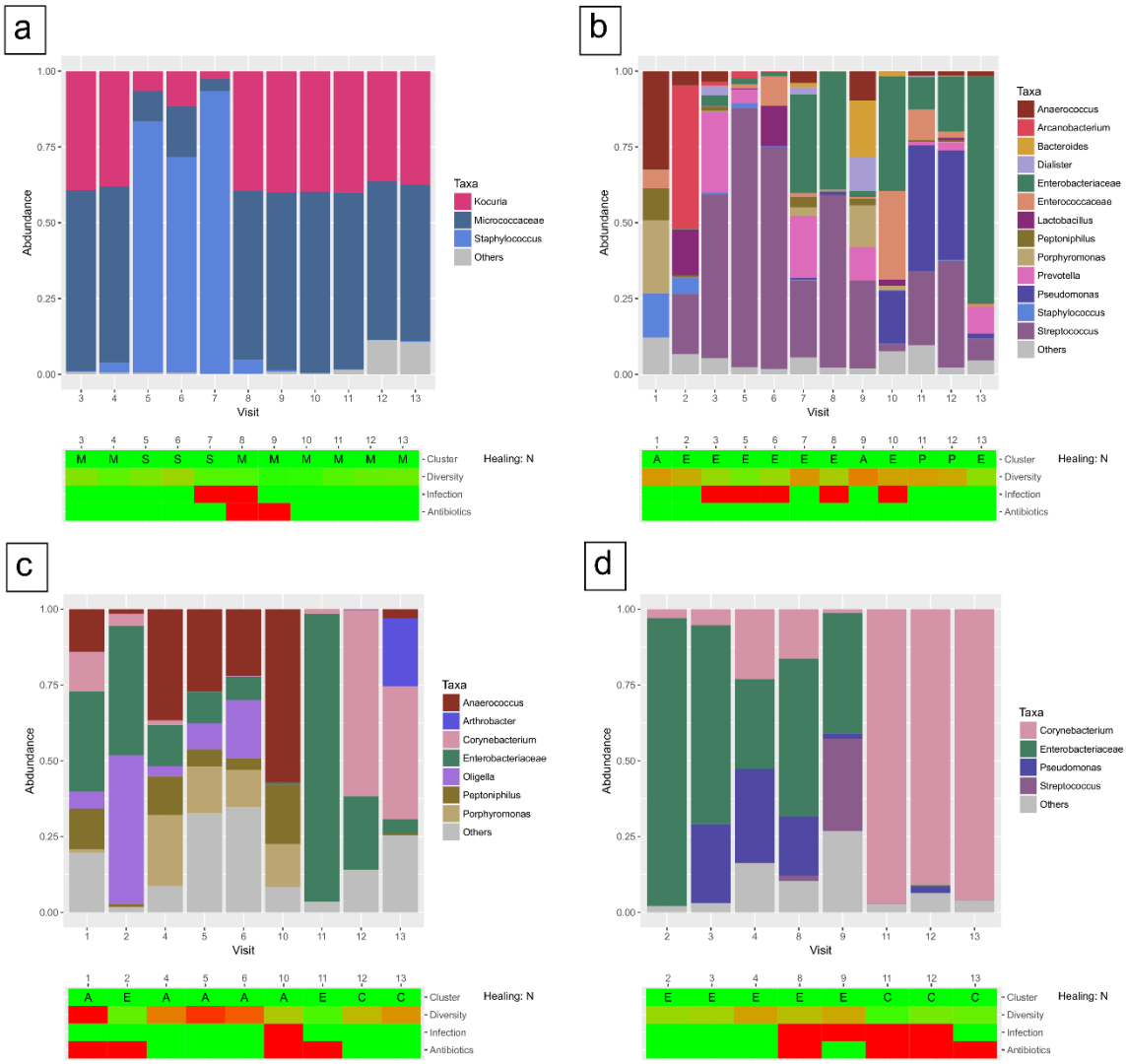
(a) Healing DFU profile largely dominated by *Corynebacterium* and *Staphylococcus*. (b) Healing DFU with transition from anaerobes and *Enterobacteriaceae* to *Corynebacterium* and *Staphylococcus*. (c) Healing DFU with disappearance of *Enterobacteriaceae* and anaerobes in the visits preceding healing. (d) Example of a non-healing DFU persistently colonised by mixed anaerobes and *Enterobacteriaceae*. Infection, antibiotic exposure or high microbial diversity is indicated by red on the colour map, while absence of infection, antibiotic exposure or low microbial diversity is indicated by green. DFU cluster designations; P = *Pseudomonaceae*, E = *Enterobacteriaceae*, C = *Corynebacteriaceae*, S = *Staphylococcaceae*, A = Anaerobes, M = *Micrococcaceae*, N = Non-healing at study end, H = Healed.


a
b
Diversity
8
6
4
2
−2
−1
0
1
2
Visits around infection onset
Visits around infection resolution
Infected
no
yes

687
688

Figure 2. Changes in diversity associated with onset and resolution of infection.

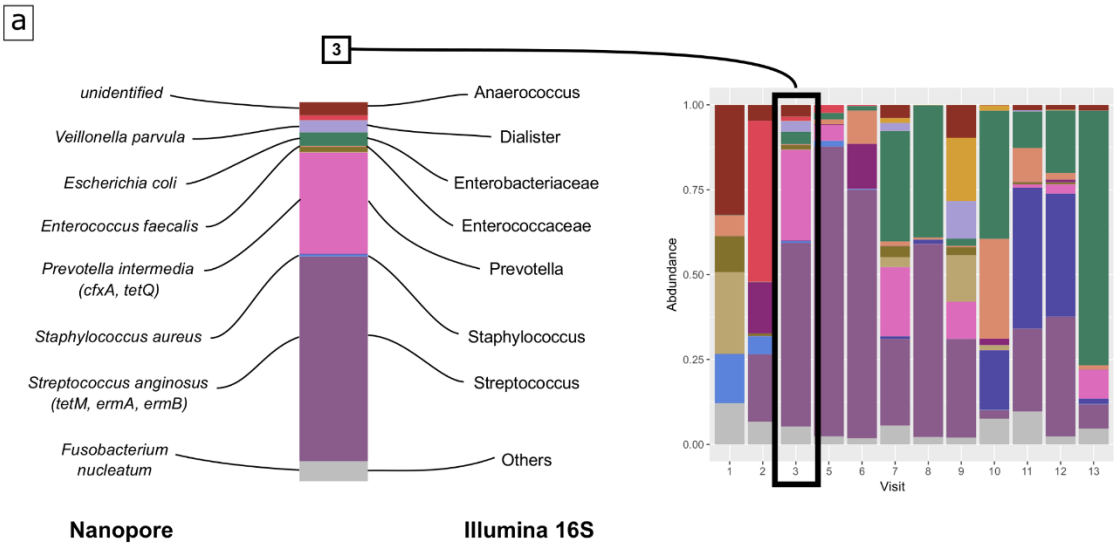
Alpha diversity (Inverse Simpson Index) plotted for DFU samples taken before, during and after all documented episodes of infection. (a) Fall in alpha diversity associated with onset of infection where visit 0 is the first visit where a new infection is observed clinically, visit -2 and -1 the two preceding visits with no infection and visit 1 and 2 any subsequent visits with persistent infection. (b) Recovery in alpha diversity associated with resolution of infection where visit 0 is the final visit where infection is observed clinically, visit -2 and -1 any preceding visits with infection and visit 1 and 2 any subsequent visits where infection had resolved. The trend line indicates changing diversity calculated by local polynomial regression fitting with the shaded area indicating the 95% confidence intervals.

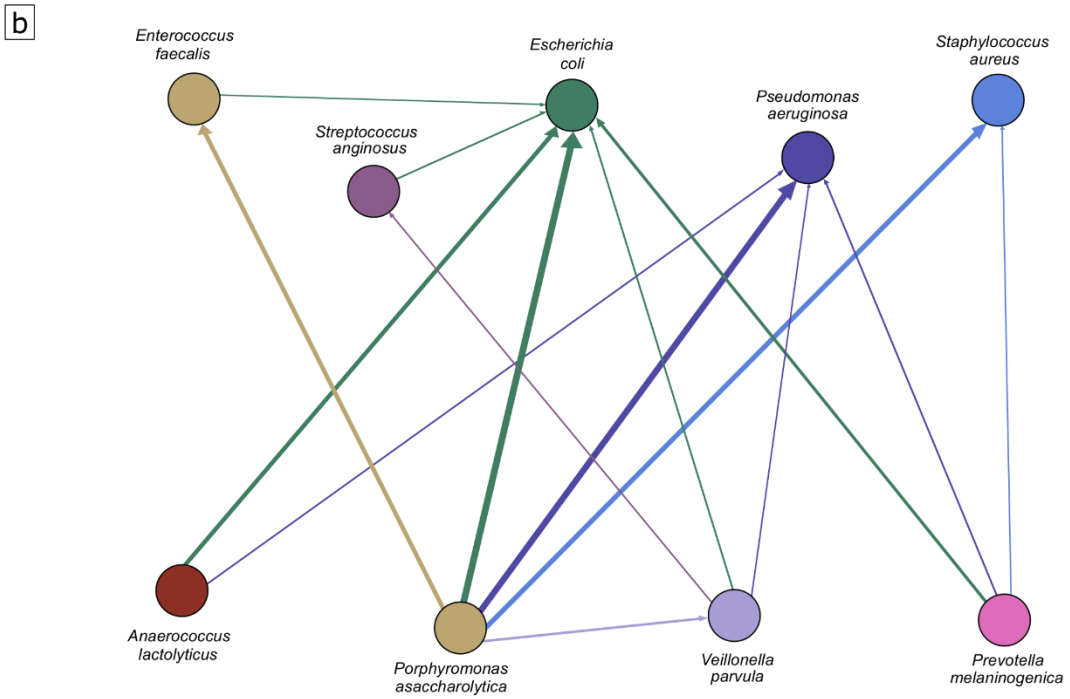

a
Abundance
Visit
Taxa
Kocuria
Micrococcaceae
Staphylococcus
Others
Cluster
Diversity
Infection
Antibiotics
Healing: N
b
Taxa
Anaerococcus
Arcanobacterium
Bacteroides
Dialister
Enterobacteriaceae
Enterococcaceae
Lactobacillus
Peptoniphilus
Porphyromonas
Prevotella
Pseudomonas
Staphylococcus
Streptococcus
Others
c
Taxa
Anaerococcus
Arthrobacter
Corynebacterium
Enterobacteriaceae
Oligella
Peptoniphilus
Porphyromonas
Others
d
Taxa
Corynebacterium
Enterobacteriaceae
Pseudomonas
Streptococcus
Others

701

Figure 3. Examples of DFU profiles where clinical infection occurred.

(a) Expansion of *Staphylococcus* associated with clinical infection at visit 7 and subsequent decline following treatment with flucloxacillin at visit 8. (b) Episodes of clinical infection observed in a profile dominated by *Streptococcus* with later emergence of *Pseudomonas* and *Enterobacteriaceae*. (c) High diversity seen at visit 6 markedly reduced at visit 10, with other anaerobes replaced by *Anaerococci* and *Peptonophilus*. An episode of clinical infection was treated with co-amoxiclav with subsequent expansion of *Enterobacteriaceae*. (d) Infection associated with emergence and expansion of *Streptococcus* followed by treatment with ciprofloxacin after visit 9 and accompanying elimination of *Streptococcus* and *Enterobacteriaceae*. Infection, antibiotic exposure or high microbial diversity is indicated by red on the colour map, while absence of infection, antibiotic exposure or low microbial diversity is indicated by green. DFU cluster designations; P = *Pseudomonaceae*, E = *Enterobacteriaceae*, C = *Corynebacteriaceae*, S = *Staphylococcaceae*, A = Anaerobes, M = *Micrococcaceae*, N = Non-healing at study end, H = Healed.


a
3
unidentified
Veillonella parvula
Escherichia coli
Enterococcus faecalis
Prevotella intermedia
(cfxA, tetQ)
Staphylococcus aureus
Streptococcus anginosus
(tetM, ermA, ermB)
Fusobacterium
nucleatum
Anaerococcus
Dialister
Enterobacteriaceae
Enterococcaceae
Prevotella
Staphylococcus
Streptococcus
Others
Abundance
1.00
0.75
0.50
0.25
0.00
1
2
3
5
6
7
8
9
10
11
12
13
Visit
Nanopore
Illumina 16S


b
Enterococcus
faecalis
Escherichia
coli
Staphylococcus
aureus
Pseudomonas
aeruginosa
Streptococcus
anginosus
Anaerococcus
lactolyticus
Porphyromonas
asaccharolytica
Veillonella
parvula
Prevotella
melaninogenica

718

719

Figure 4. Direct sample sequencing enables identification of species and resistance genes

(a) Detailed analysis of the DFU profile shown in Fig. 3(b), comparing taxonomic identifications based on Illumina 16S rRNA and Oxford Nanopore MinION direct sample sequencing taken during an episode of clinical infection at visit 3.

(b) Potential network of inter-species interaction for representative species based on this DFU profile, created using the Metabolic Complementary Index (23). Edges are shown for all scores greater than 0.5. Higher scores are reflected by greater edge thickness with arrows directed towards the partner species from which metabolic benefit is predicted to derive.

Supplementary Table 1. Patient demographics and DFU details.

Age	Sex	DFU depth (NPUAP)	DFU size (mm^2)	Healed	Infection	Alpha diversity (ISI)	Profile figure reference
37	M	II	70	Yes	No	3.25	
63	M	II	261	No	No	2.05	
62	M	III	618	No	Yes	3.53	
84	M	III	30	No	Yes	5.25	
60	M	III	1203	No	Yes	5.73	Fig. 3(c)
58	M	III	50	No	No	1.60	
77	F	III	251	No	No	5.22	
74	M	III	103	No	No	1.44	
45	M	III	1713	No	Yes	3.21	Supp. Fig. 4(c)
68	M	III	700	Yes	No	2.33	Fig. 1(b)
73	M	II	891	No	No	3.03	Fig. 1(d)
71	M	II	19	No	No	3.15	Supp. Fig. 4(a)
42	F	II	109	Yes	No	1.04	Fig. 1(c)
81	F	II	930	No	Yes	1.11	Fig. 3(d)
71	M	III	344	No	Yes	2.93	
61	M	III	146	Yes	Yes	3.30	Fig. 1(a)
79	F	III	287	No	Yes	1.10	Supp. Fig. 4(d)
81	F	III	114	No	Yes	1.96	Fig. 3(a)
56	M	II	107	No	No	3.28	
88	M	II	61	No	No	3.22	
62	M	III	25	Yes	Yes	2.85	
70	M	III	37	No	Yes	3.03	
56	M	III	33	No	No	1.05	
76	F	III	177	No	No	1.09	Supp. Fig. 4(b)
73	M	IV	990	Yes	No	1.74	
79	M	III	69	Yes	Yes	1.10	
50	F	III	593	No	Yes	4.79	Fig. 3(b), Fig. 4
85	F	II	184	Yes	No	2.49	

Baseline patient demographics and DFU details for all 28 study participants. DFUs which healed or developed clinical infection during the study are indicated by 'Yes' or 'No' in the Healed and Infection columns respectively. NPUAP = National Pressure Ulcer Advisory Panel depth grade. ISI = Inverse Simpson Index.

Supplementary Table 2. Highly abundant taxa.

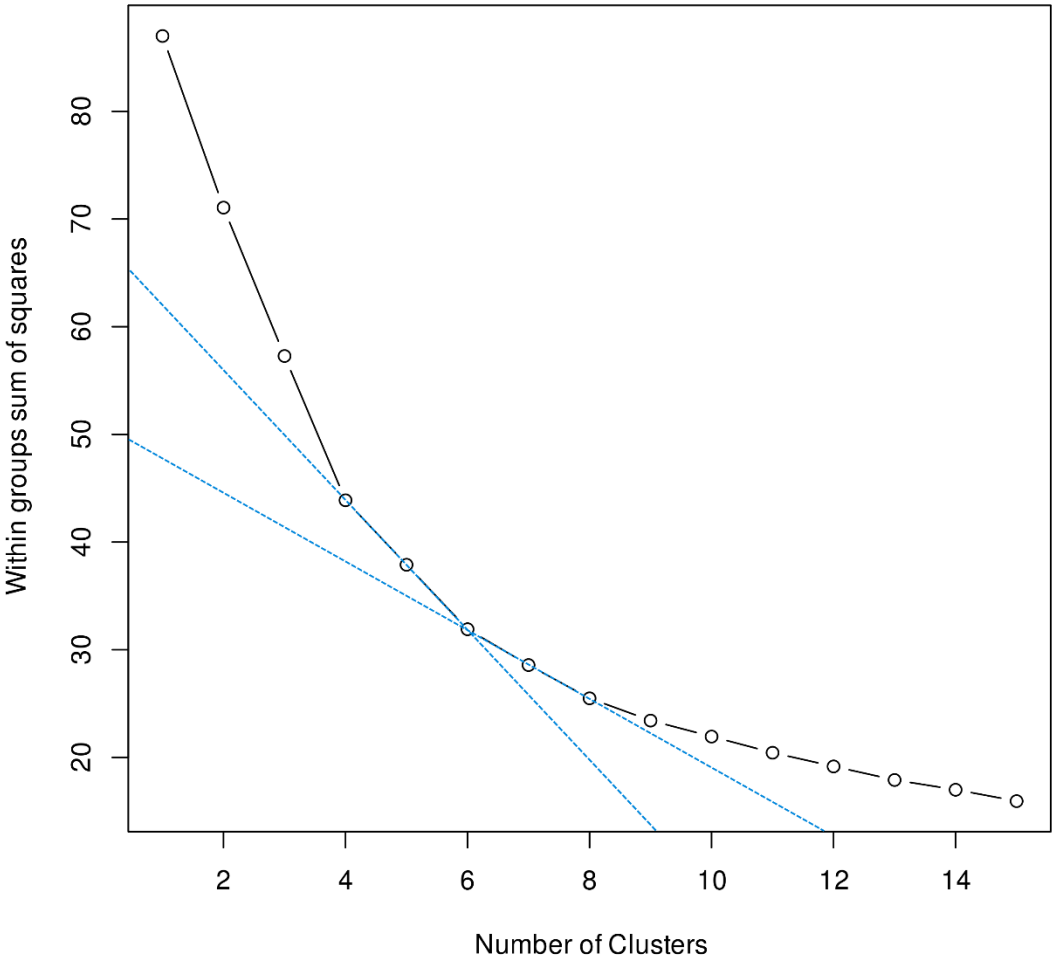
Family	**Abundance (%)**	**Dominant genera**
Corynebacteriaceae	22.7	*Corynebacterium* (22.7 %)
Staphylococcaceae	15.3	*Staphylococcus* (15.2 %)
Clostridiales Incertae Sedis X	13.4	*Anaerococcus* (6.4 %), *Finegoldia* (2.4 %), *Helcococcus* (1.8 %)
Enterobacteriaceae	11.5	Unclassified (10.6 %), *Morganella* (0.5 %)
Micrococcaceae	7.9	*Kocuria* (3.9 %), Unclassified (3.1 %), *Arthrobacter* (0.8 %)
Pseudomonadaceae	6.1	*Pseudomonas* (6.1 %)
Streptococcaceae	4.6	*Streptococcus* (4.6 %)
Moraxellaceae	3.2	*Acinetobacter* (3.0 %)
Prevotellaceae	2.8	*Prevotella* (2.8 %)
Actinomycetaceae	2.1	*Arcanobacterium* (0.9 %), *Actinobaculum* (0.8 %)

The 10 most abundant taxa at family level are shown according to mean proportional abundance across all 237 DFU samples. Dominant genera are indicated for each family.

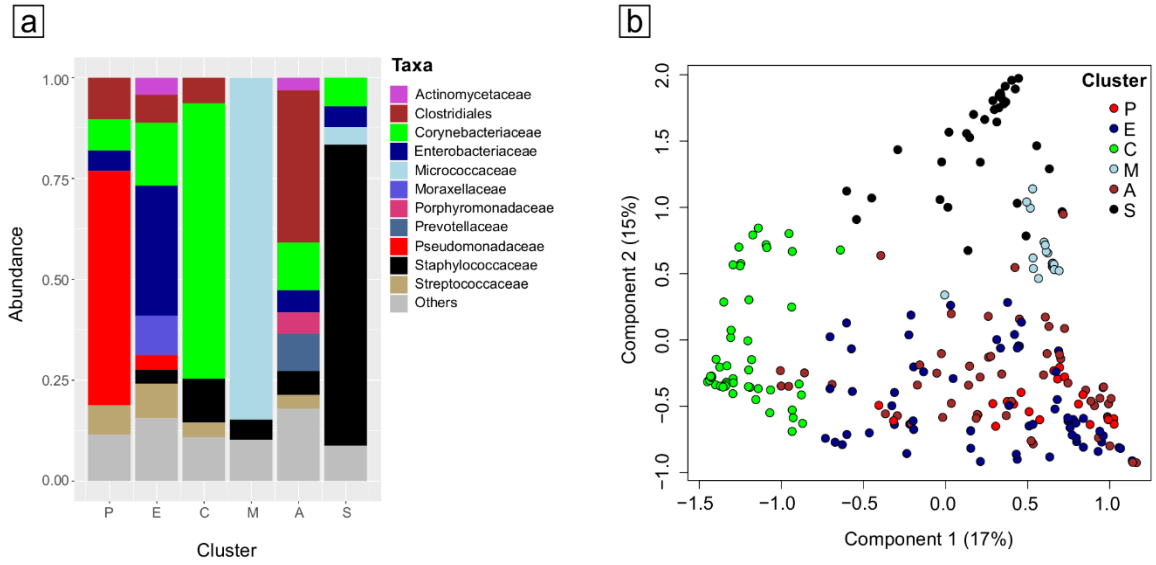
Supplementary Table 3. DFU cluster transition probabilities

All DFUs									
From	**To**								**Count**
	P	**E**	**C**	**M**	**A**	**S**	**N**	**H**	
P	*0.65*	0.15	0.05	0.00	0.05	0.00	0.05	0.05	20
E	0.05	*0.56*	0.11	0.00	0.08	0.08	0.10	0.02	61
C	0.02	0.06	*0.67*	0.00	0.06	0.02	0.10	0.06	49
M	0.00	0.00	0.00	*0.71*	0.00	0.18	0.12	0.00	17
A	0.02	0.20	0.04	0.00	*0.64*	0.04	0.07	0.00	56
S	0.00	0.09	0.06	0.09	0.09	*0.53*	0.06	0.09	34
Healed DFUs									
From	**To**								**Count**
	P	**E**	**C**	**M**	**A**	**S**	**N**	**H**	
P	*0.50*	0.17	0.00	0.00	0.17	0.00	0.00	0.17	6
E	0.11	*0.11*	0.22	0.00	0.11	0.33	0.00	0.11	9
C	0.00	0.00	*0.65*	0.00	0.12	0.06	0.00	0.18	17
M	0.00	0.00	0.00	*0.00*	0.00	0.00	0.00	0.00	0
A	0.13	0.25	0.25	0.00	*0.38*	0.00	0.00	0.00	8
S	0.00	0.09	0.09	0.00	0.00	*0.55*	0.00	0.27	11
Unhealed DFUs									
From	**To**								**Count**
	P	**E**	**C**	**M**	**A**	**S**	**N**	**H**	
P	*0.71*	0.14	0.07	0.00	0.00	0.00	0.07	0.00	14
E	0.04	*0.63*	0.10	0.00	0.08	0.04	0.12	0.00	52
C	0.03	0.09	*0.69*	0.00	0.03	0.00	0.16	0.00	32
M	0.00	0.00	0.00	*0.71*	0.00	0.18	0.12	0.00	17
A	0.00	0.19	0.00	0.00	*0.69*	0.04	0.08	0.00	48
S	0.00	0.09	0.04	0.13	0.13	*0.52*	0.09	0.00	23

Transition probabilities for movement between DFU clusters over subsequent visits for all DFUs (n = 28) and also separated by DFUs which went on to heal (n = 8) and those that did not (n = 20). Probability of movement between clusters is shown from 0 (no probability) to 1 (guaranteed probability) with the number of samples falling into each of the 6 clusters shown in the count column. Self-transitions are italicised. P = *Pseudomonaceae*, E = *Enterobacteriaceae*, C = *Corynebacteriaceae*, M = *Micrococcaceae*, A = Anaerobes, S = *Staphylococcaceae*, N = non-healing state (study end), H = healed DFU.

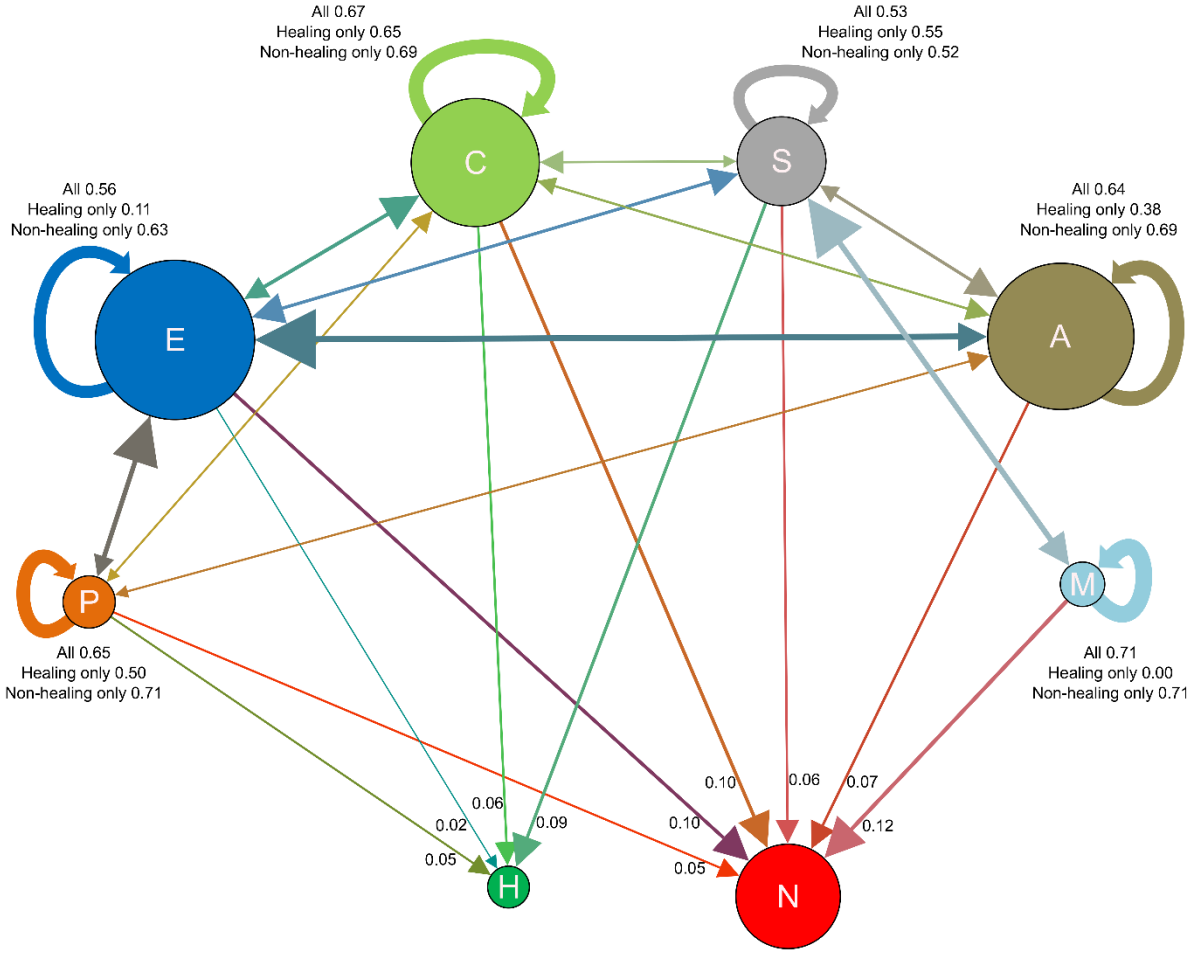


Supplementary Figure 1. Cluster variance explained by increasing number of clusters. Plot of variance explained by increasing the numbers of clusters used, showing the trade-off between clustering complexity and fit to the data. The change in trajectory in moving from 5 to 6 clusters compared to 6 to 7 clusters is shown with dashed lines indicating the diminishing returns associated with using a higher number of clusters.



Supplementary Figure 2. Clustering of DFU samples by major taxonomic groups.

DFU samples were clustered at family level by k-means clustering and assigned to one of 6 clusters, labelled according to most abundant taxa. DFU cluster designations; P = *Pseudomonaceae*, E = *Enterobacteriaceae*, C = *Corynebacteriaceae*, S = *Staphylococcaceae*, A = Anaerobes, M = *Micrococcaceae*. (a) Mean proportional abundance of the taxa in each cluster. The E and A clusters contained were the most heterogeneous, representing a more diverse combination of different taxa. (b) Principal coordinates analysis showing the distribution of samples at genus level, with samples coloured by cluster. Samples in the low diversity C, M and S clusters were well discriminated by the first two principal components while samples from clusters E, A and P tended to overlap and were more widely spread. DFU clusters accounted for a significant proportion of the overall variance when analysed with PERMANOVA (7 % variance explained, p = 0.001).

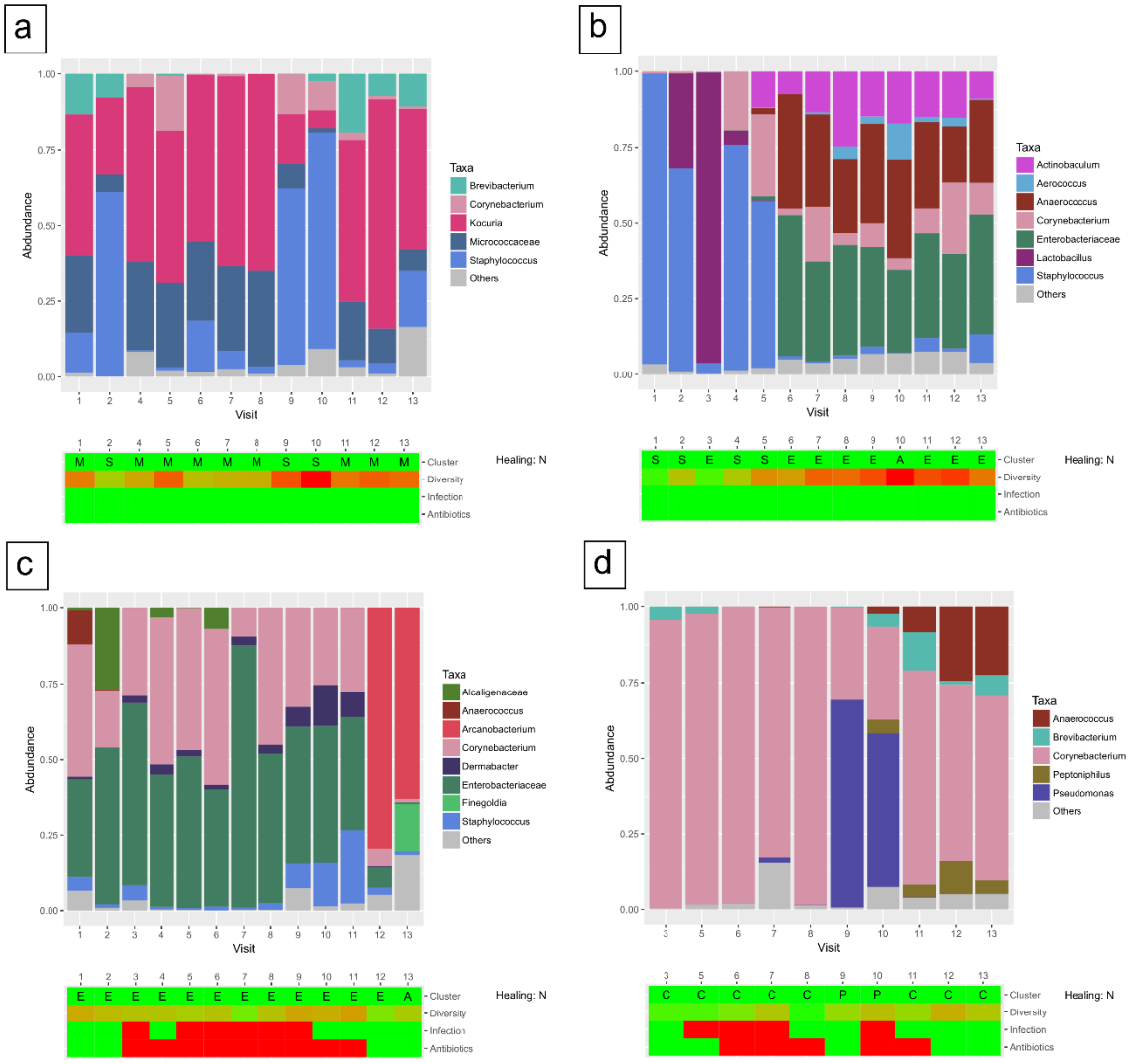


778
779

780 **Supplementary Figure 3. Transition between DFU clusters during the study.**

781 A Markov model demonstrating the transition between DFU clusters between visits
782 until either healing (H) or non-healing at study end (N) for 237 samples from 28
783 subjects combining both healing (n=8) and non-healing (n=20) DFUs. The size of the
784 nodes indicates the number of samples in each state while the edge and arrow weights
785 represent the probability of transition between states. Transition probabilities within
786 nodes or to the endpoints (healing or non-healing) are labelled, with self-transition
787 probabilities separated by group also shown. The full transition data are also shown
788 in Supplementary Table 3. DFU cluster designations; P = *Pseudomonaceae*, E =
789 *Enterobacteriaceae*, C = *Corynebacteriaceae*, S = *Staphylococcaceae*, A =
790 Anaerobes, M = *Micrococcaceae*.

791



Supplementary Figure 4. Additional examples of DFU profiles.
Examples of DFU profiles with unexplained abrupt changes in microbial profile, or static profiles despite clinical infection. (a) Proliferation of *Staphylococcus* at visits 2 and 9 without clinically apparent infection, (b) Transition from colonisation with *Staphylococcus* to *Enterobacteriaceae* and anaerobes at visit 6, (c) Stable colonisation with *Enterobacteriaceae* and *Corynebacterium* despite clinical infection and antibiotic exposure, (d) Colonisation with *Corynebacterium* unchanged by infection or antibiotic exposure, with a second episode of infection associated with proliferation of *Pseudomonas.* Infection, antibiotic exposure or high microbial diversity is indicated by red on the colour map, while absence of infection, antibiotic exposure or low microbial diversity is indicated by green. DFU cluster designations; P = *Pseudomonaceae*, E = *Enterobacteriaceae*, C = *Corynebacteriaceae*, S = *Staphylococcaceae*, A = Anaerobes, M = *Micrococcaceae*, N = Non-healing at study end, H = Healed.