

SCIENTIFIC DATA

OPEN

DATA DESCRIPTOR

16S rDNA based skin microbiome data of healthy individuals and leprosy patients from India

Nitin Bayal¹, Sunil Nagpal², Mohammed Monzoorul Haque², Milind S. Patole¹, Vijayalakshmi Valluri³, Rohini Suryavanshi⁴, Sharmila S. Mande² & Shekhar C. Mande^{5*}

Leprosy is an infectious disease that has predilection in skin and peripheral nerves. Skin has its own microbiome, however it is not extensively studied in Indian leprosy patients. Here, by using next-generation 16S rDNA sequencing, we have attempted to assess the skin associated microbial diversity pertaining to affected and unaffected skin of Indian leprosy patients. A total of 90 skin swab samples were collected from 60 individuals (30 healthy controls, 30 patients) residing in Hyderabad and Miraj, two distinct geographical locations in India to assess the homo/heterogeneity of skin microbial signatures. While a large increase in genus *Methylobacterium* and *Pseudomonas* was seen in patients from Miraj and Hyderabad respectively, a considerable decrease in genus *Staphylococcus* in the leprosy patients (as compared to controls) from both geographical locations was also observed. We expect that, these datasets can not-only provide further interesting insights, but will also help to observe dynamics of microbiome in the diseased state and generate hypotheses to test for skin microbiome transplantation studies in leprosy.

Background and Summary

Microbiome and dermatology research are related to each other¹⁻³. Microorganisms generally thrive as site-specific close knit communities, however, at species and strain level, they can be as unique as human fingerprints¹. Skin is the largest organ of human body and one of the most complex and unexplored areas in microbiome investigation²⁻⁶. High-throughput sequencing technologies have become the main stay of modern scientific studies in microbiome space^{7,8}. Particularly, 16S ribosomal RNA gene analysis is considered as a gold standard for taxonomic identification of bacteria, also referred to as meta-taxonomic analysis^{9,10}.

The morphology and physiological characteristics of various sites on the human skin define the composition of microbial communities^{11,12}. Besides serving as a physical barrier that resists infiltration by potential pathogenic bacteria, skin microbiome is known to play diverse physiological roles in maturation and homeostasis of cutaneous immunity^{13,14}. In cutaneous biology, either a shift in microbiome, genetic predisposition, misuse of antibiotics or loss of beneficial organisms is believed to aggravate the inflammatory skin disease conditions¹⁵. This is known as dysbiosis, which leads to an aberrant activation of immune system, resulting in compromised regulation of immune responses to commensal microbes¹⁶. There are various dermatological conditions that have visual and technical differences in their symptoms and diagnosis^{17,18}.

Leprosy is one of the serious skin disorders found to have been linked to dysbiosis¹⁹. It is a chronic systemic granulomatous disease that primarily affects the skin and peripheral nerves, and less commonly mucosa of the upper respiratory tract, testes and the anterior chamber of the eyes²⁰. It is caused by bacteria namely, *Mycobacterium leprae*²¹ and *Mycobacterium lepromatosis*^{22,23}. If left untreated, leprosy can cause irreversible damage²³ to its manifested body parts, which is one of the reasons for the attached social stigma^{24,25}. Ridley and Jopling classified leprosy in a variety of clinical forms viz. tuberculoid (TT/ paucibacillary), borderline (BL), or lepromatous (i.e. multibacillary LL/MB) based on the immune response to the organism²⁶.

¹National Centre for Cell Science, NCCS Complex, Pune, India. ²BioSciences R&D, TCS Research, Tata Consultancy Services, Pune, India. ³Immunology & Molecular Biology Department, Bhagwan Mahavir Medical Research Centre, A. C. Guards, Hyderabad, Telangana State, India. ⁴Richardson Leprosy Hospital (The Leprosy Mission Trust), Miraj, Maharashtra, India. ⁵Council of Scientific & Industrial Research, Anusandhan Bhawan, 2, Rafi Marg, New Delhi, India. *email: shekhar@csir.res.in

Studies on Acne^{27,28}, Vitiligo²⁹, Atopic dermatitis^{30,31} and Psoriasis³² have shown that these dermatological conditions exert influence on commensal microbiome and exhibit compositional variation in the disease onset or progression at the lesional site and at the systemic level. The extent to which skin microbiota changes after a systemic antibiotic intervention depends on the chemical nature of the antibiotic, combinations of antibiotics, type of administration, duration and dose. Although skin is physically compartmentalized from other body sites, cross-inoculation can lead to spread of antibiotic resistant microorganisms to skin³³.

A key study involving Brazilian leprosy patients provides an in-depth analysis of the composition of skin microbiome in lepromatous, skin lesions (and paired adjoining non-lesional areas) sampled from a Brazilian patient cohort³⁴. Results of this study primarily indicate that both lesional as well as non-lesional skin of (treated/untreated) individuals infected with leprosy harbors microbial communities that are significantly less diverse as compared to the skin of healthy individuals. This reduced diversity could possibly be attributed to the impact of the infectious agent (*M. leprae*) itself or it could be a systemic change resulting due to the ongoing MDT regime. Furthermore, there is a strong association between the lung microbiome and tubercle formation in an infection caused by *Mycobacterium tuberculosis*³⁵. Similarly whether a similar relationship exists between skin microbiome and progression of leprosy needs further studies.

It is of interest to study and contemplate the relationship between dysbiosis of skin microbiota and its effect on host skin metabolism and outcome of the leprosy. Given this context, we endeavored to study the structure of skin microbiome in healthy controls as well as of leprosy patients from India, a country that accounts for the highest annual number of leprosy cases diagnosed/reported worldwide. In 2016, 34,000 new cases were reported alone from endemic states in India with 25% annual cases³⁶. Given high variability with respect to dietary habits, climatic and life-style conditions within India, we chose two well-separated geographical locations within India (viz. Lepra Blue Peter Public Health and Research Centre, Hyderabad and TLM Richardson Leprosy Hospital, Miraj) to check if there are differences created as a result of such potential confounding factors. These centres are specialized in diagnosis, treatment and management of leprosy including studies in related epidemiology.

This is the first study of dysbiosis of skin microbiome in leprosy patients as compared to healthy individuals from India. We report data from 60 study participants including 30 healthy individuals and 30 patients who were undergoing oral multi-drug treatment (MDT) therapy. The subsequent sections of this report provide details pertaining to study design, sample collection methodology, types of metadata collected, bioinformatics methods used for obtaining taxonomic abundance tables corresponding to sequence data, and the types of data that are being provided along with this report for public use.

Methods

Study design. Figure 1 depicts the overall design of the present study. The present study was primarily an observational; non-interventional study conducted at two distinct centers viz. Hyderabad and Miraj (i.e. two Indian cities approximately 500 kilometers apart). The Institutional Ethical Committees of National Centre for Cell Science (NCCS), Pune as well as the respective sampling locations approved the study (IEC Approval Ref. No. NCCS/IEC/2016-I/1). The collection protocol was developed in line with the Nuremberg Code³⁷ and Declaration of Helsinki³⁸. Personal identifiable information pertaining to all participants was removed to ensure confidentiality^{39,40}. All participants were provided a detailed explanation of study objectives and the sampling procedure, and written consent documents were obtained. Cohort size for the study was not statistically derived. The following inclusion and exclusion criteria were used for selection of consenting study participants (who were diagnosed for leprosy disease).

Inclusion criteria.

1. Age between 21 to 70 years.
2. Patients having a high (3+) bacillary index (MB/BL/LL).
3. Active patches present on the patients back (scapular and lumbar regions).

Exclusion criteria.

1. Pregnant women.
2. Patients co-infected with Leprosy-TB or Leprosy-HIV.
3. Individuals with a different illness or those who have taken any medications in previous 10 days for any other symptoms.

Sample collection process. Person collecting samples mandatorily wore a sterilized clean lab apron, double layered gloves, hair cap and laboratory facemask to avoid any sort of cross-contamination. Sampling was done by researchers in a dedicated sample collection room. A fresh pair of gloves was used while collecting samples from every new participant of this study. As the aim of the study was to study skin microbiome, a non-invasive method of collection of skin swabs was performed instead of skin scrapings or skin biopsy^{41,42}. Individual's comfort was also taken in priority while sampling. A swabbing procedure was employed wherein two samples (5 cm × 5 cm) were collected from an affected lesional area and an adjoining uninfected portion (5–8 cms away from the lesional site). The skin swabs were collected from back area of the body using HiCulture Sterile Skin Swab collection device (HiMedia Labs, India) soaked in wetting solution (0.15 M NaCl with 0.1% Tween 20). The swabbing technique was executed as follows: Skin areas were selected and stretched with one hand, the swab in the other hand was held in such a manner that that the swab shaft is parallel to skin surface. Subsequently, the swab was moved in one direction for fifty times, thrice facing each side of the swab applying firm pressure. The collected swab samples were stored in –80 °C until further processing. Section A of Data Record titled, 'Sample

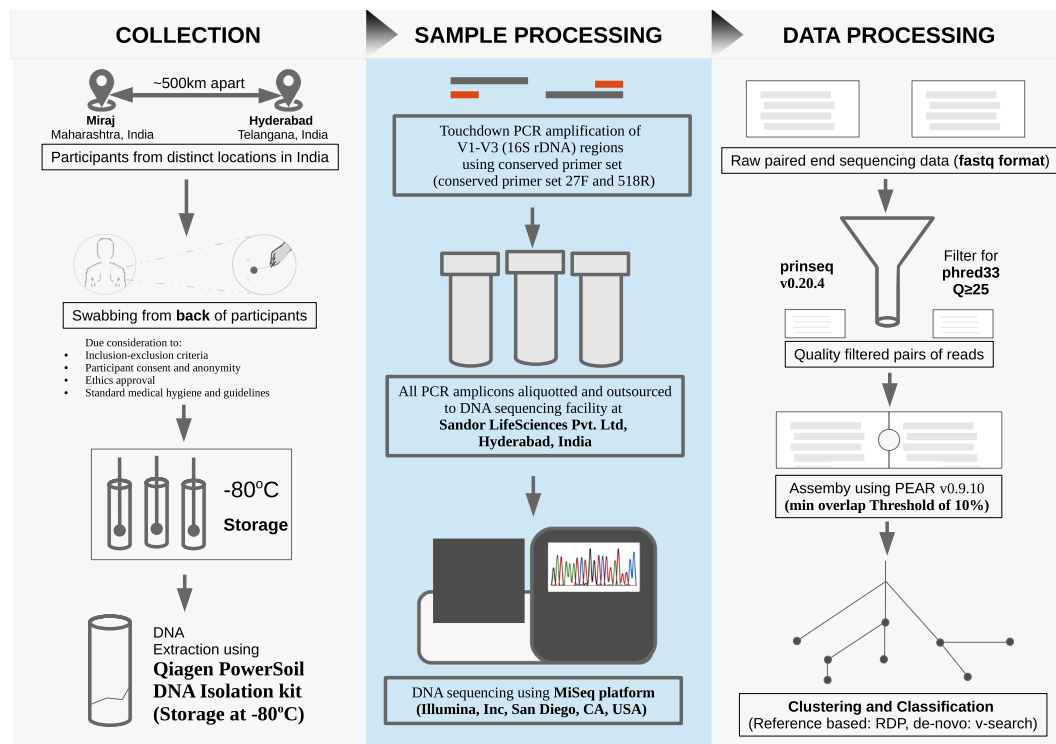


Fig. 1 Schematic workflow depicting overall study design and methodology followed for sample collection, processing and data processing.

Collection Image' depicts an image of protocol followed for skin swab sampling and Section B represents an image for a lesional site of one of a leprosy patient suffering from lepromatous leprosy as per records.

DNA isolation and processing. DNA extraction was carried out using Qiagen PowerSoil DNA Isolation kit protocol with indicated modifications. Each swab was cut and swirled in a 2 ml collection tube containing MicroBead Solution (MO BIO Laboratories, Carlsbad, CA). The swab sponge was pressed against the wall of tube multiple times to ensure transfer of bacteria from the swab to the solution. Subsequently, swabs were given temperature shock twice at 90 °C in a water bath for 20 minutes and then at –80 °C to kill the bacteria. DNA purity was accessed by absorbance at 260 nm using a NanoDrop 1000 Spectrophotometer (Thermo Fischer Scientific, USA). The extracted DNA was stored at –20 °C until further processing.

V1-V3 hyper-variable regions PCR amplification and sequencing. Amplification of the V1-V3 hyper-variable regions was performed using the bacterial 16S rRNA gene region specific primers. It was amplified using touchdown PCR with the conserved primer set 27F (5'–AGAGTTTGATCCTGGCTCAG–3') and 518R (5'–GTATTACCGCGGCTGCTGG–3'). The amplicons were gel purified and then quantified by absorbance at 260 nm using a NanoDrop 1000 Spectrophotometer (Thermo Fischer Scientific, USA). DNA purity was assessed using the A260/A280 ratio. Samples failing amplification were amplified again using fresh aliquot until a satisfactory PCR product (for Illumina sequencing) was yielded. All the PCR amplicons were aliquoted and outsourced to DNA sequencing facility at Sandor Lifesciences Pvt. Ltd, Hyderabad, India. Sequencing was performed using the MiSeq platform (Illumina, Inc, San Diego, CA, USA) according to manufacturer's instructions.

Bioinformatics analysis. Total samples collected in this study were 90, however given the read quality issues with one of the unaffected region samples (Sample A9), paired exclusion was performed for Sample A9 and its associated sample from the same participant i.e. Sample G7. Consequently, a total of 88 skin swab samples corresponding to (29) leprosy subjects and (30) controls (from Hyderabad and Miraj region of India) were used for the analysis. Table 1 provides an overview of the number of subjects and samples collected from the two mentioned research locations.

Data records file titled, 'Metadata for Leprosy Microbiome Study' provides a comprehensive metadata information table corresponding to all collected samples.

Pre-processing. Raw reads of paired-end sequence data was subjected to standardized quality processing. High quality pairs with a minimum phred33 quality score of 25 were retained using prinseq v0.20.4⁴³. Assembly of high quality pairs was performed using PEAR v0.9.10⁴⁴ with a minimum overlap threshold of 10%. Only those reads were retained which had a minimum merged length of 400 bp. As mentioned above, given the poor quality of reads, one sample (Sample ID: A9 from an unaffected area of a study participant having leprosy) and

Participant Type	Hyderabad		Miraj		Total Samples
Healthy (Control)	15		15		30
Patients	Affected (Lesions)	Unaffected	Affected (Lesions)	Unaffected	60
	14	14	16	16	

Table 1. An overview of the number of study participants and samples collected from the two-research locations viz. Hyderabad and Miraj.

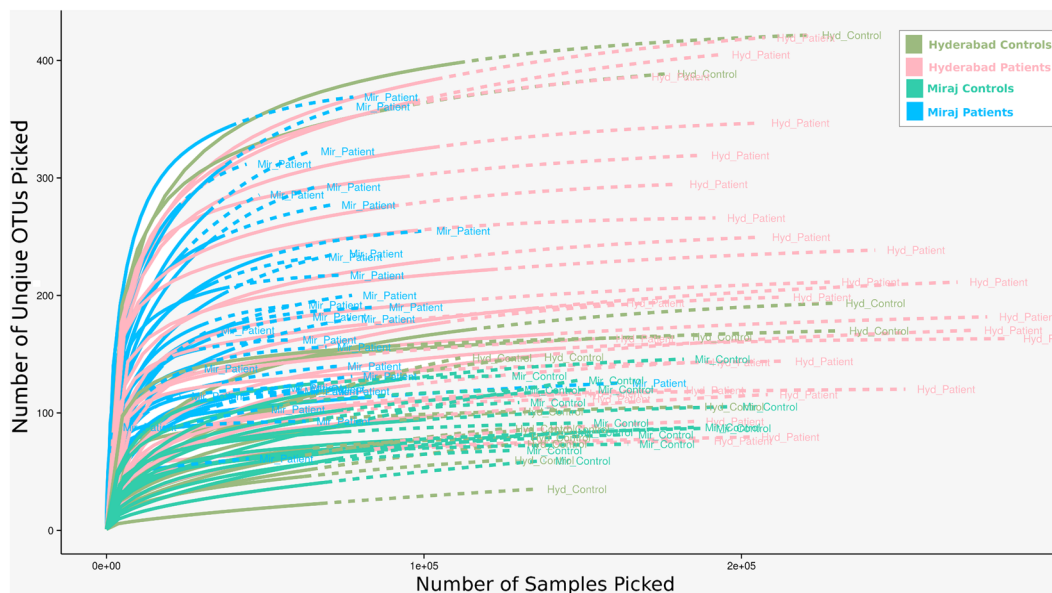


Fig. 2 Rarefaction curves indicating sequencing depth for the samples collected in this study.

its corresponding sample (taken from the affected region (G7) of the same participant) was discarded during the analysis process.

Taxonomic classification. The quality checked and assembled sequence data was then subjected to closed reference classification using RDP classifier v.2.12⁴⁵ at 0.8 minimum assignment confidence thresholds. Abundance data matrices at all five major levels of taxonomic lineage i.e. Phylum, Class, Order, Family and Genus were compiled using the hier files of RDP classification output (See section: Data Records). The sequence data was also processed using a de-novo OTU classification approach at a clustering identity of 0.97. The program vsearch v2.8.0_linux_x86_64⁴⁶ was used as per the default protocol mentioned (for chimera removal and OTU clustering) at <https://github.com/torognes/vsearch/wiki/VSEARCH-pipeline>. Chimera removal was done using both protocols viz. mapping against the Gold reference database as well as the *de novo* chimera detection steps mentioned in the vsearch pipeline.

Rarefaction. Rarefaction curves were plotted for the genus level abundance data using the observed species index generated by iNEXT v2.0.9⁴⁷ (at 50 bootstrap replications for a knot size of 40 and end point equal to double the sample size) and ggplot2 v2.1.0.1⁴⁸ R packages. Except for a few samples from Miraj, results of rarefaction depicted in Fig. 2 indicated sufficient sequencing depth for majority of the analyzed samples. The range of sample sizes for RDP assigned genus level OTUs was observed to be 543–1,31,889, with minimum sample size for Sample B2_16B and maximum size for G8_S15.

Preliminary taxonomic assessment. Raw RDP taxonomic profiles generated indicated the presence of 1016 unique genera corresponding to 235 families, 90 orders, 55 classes and 25 phyla. Similarly, de novo OTU clustering generated a total of 161,695 OTUs. Two taxonomic abundance tables were generated (Refer section: Data Records). While the first table comprises taxonomic abundances of all 161,695 OTUs, the second table contains classification results for only those OTUs (28,321), which had a cumulative raw abundance value of two or more in the entire sample set (i.e. all population level singleton OTUs were discarded). Since the latter table has been filtered as per the mentioned criterion, we have provided the raw abundance table as well to provide researchers the flexibility to filter the same using any other criteria of their choice (Refer section: Data Records).

Figure 3a,b depict the relative abundance of top (RDP classified) genera in samples from study participants belonging to Hyderabad and Miraj respectively. Figure 4a,b depict the relative abundance of top (RDP classified) de novo generated OTUs in samples from study participants belonging to Hyderabad and Miraj respectively.

Preliminary analysis of taxonomic trends indicates the following three interesting trends.

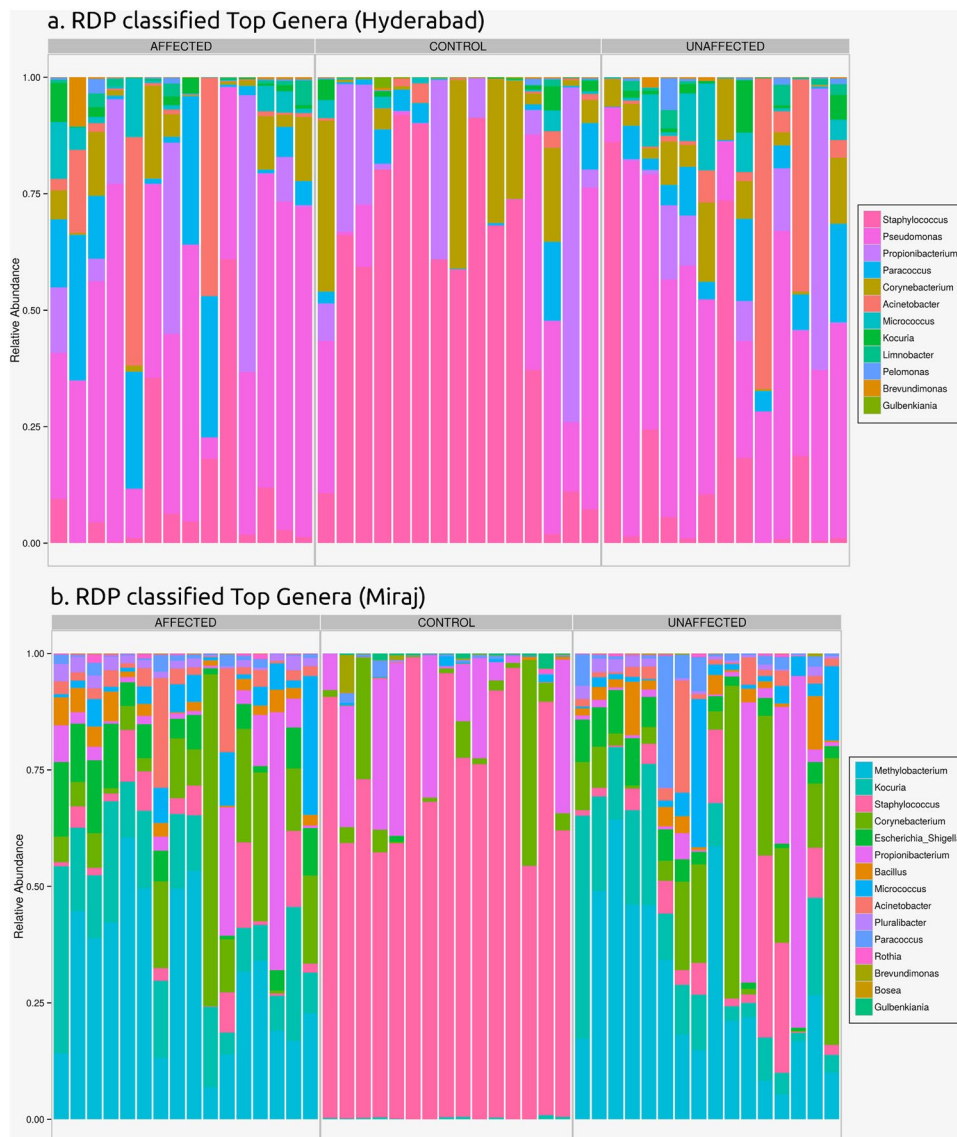


Fig. 3 Relative abundance of top (RDP classified) genera in samples from study participants belonging to Hyderabad and Miraj locations.

- (1) In both sets of samples (belonging to Hyderabad and Miraj), a stark difference is observed between the taxonomic profiles corresponding to skin microbiome samples of healthy controls as compared to that from participants affected with leprosy. Figures indicate a distinct depletion of *Staphylococcus* in samples taken from leprosy affected participants.
- (2) There appears to be uniformity in skin microbiome profiles of healthy controls irrespective of geographical location (Hyderabad and Miraj). In contrast, the skin microbiome profiles of samples taken from leprosy affected participants from Hyderabad appear to have significant differences from that taken from affected participants from Miraj.
- (3) Absence of a distinct difference between samples taken directly from the leprosy lesional sites and those taken from adjoining unaffected regions.
- (4) Data record titled 'Dendrogram Embedded Heatmap' provides an overview of the relative compositional profile of skin microbiome across various classes of samples collected in this study. The heatmap was obtained by using Microbiomeanalyst⁴⁹ at a prevalence threshold of one hit in at-least 20% of samples of the entire dataset. A centered log ratio transformation (CLR) of data was performed for the generation of heatmap. Euclidean distance was employed for generating OTU dendrogram embedded in the plot.

Ordination. In order to obtain preliminary insights into the compositional similarities (and differences) between samples pertaining to controls and patients, as well as the impact of different geographical affiliations, Jansen-Shannon divergence (JSD) and Partitioning Around Medoids (PAM) based PCoA was employed^{27,50}. Figure 5 represents the ordination analysis performed using all samples in this study with meta-data affiliation pertaining to both state of health of participants and geography of sampling. It was interesting to observe

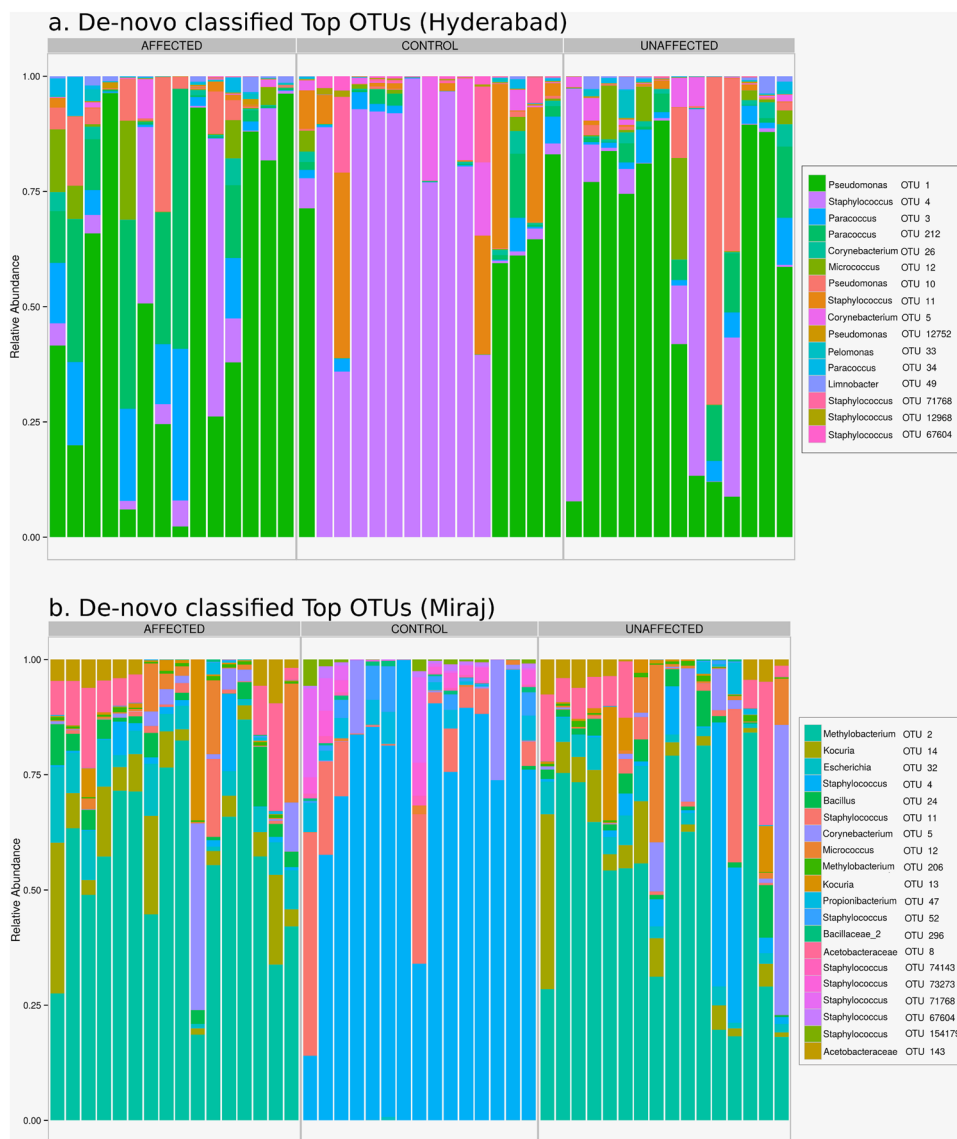


Fig. 4 Relative abundance of top (RDP classified) de-novo generated OTUs (using vsearch) in samples from study participants belonging to Hyderabad and Miraj locations.

clustering together of control samples irrespective of their geographical affiliation. In addition, geography specific segregation of samples pertaining to patients from Miraj and Hyderabad centers was apparent. Data record titled, 'Dendrogram Embedded Heatmap' (as introduced earlier) further depicts a dendrogram embedded heatmap based on taxa abundances in each of the samples of the entire study. Given the preliminary nature of these observations, a detailed analysis of the depicted pattern and assessment of microbiome characteristics with respect to its community diversity (alpha and beta) characteristics, associations or correlations between microbial abundances and associated metadata, inferred functional analysis using tools like Picrust⁵¹ and/or iVikodak⁵² is beyond the scope of the current article. However, we anticipate that the data we have provided along with this report will allow researchers to investigate and compare the provided skin microbiome profiles with other available analogous datasets.

Data Records

We share three data types corresponding to the present study. We have made sincere attempts to organize the data in a manner that makes it amenable for further analysis. While Data type 1 (i.e 16S rDNA sequence data) is available for download at the NCBI Sequence Read Archive (Study: SRP187334⁵³), the other two data types (i.e. Metadata and Taxonomic abundance datasets⁵⁴) have been provided as a part of this publication (accessible through figshare).

FASTQ data. The sequencing data corresponding to all 90 samples are accessible at NCBI Sequence Read Archive (SRA) as BioProject: PRJNA505133 (Study: SRP187334⁵³); Title: 16S rDNA based skin bacterial flora

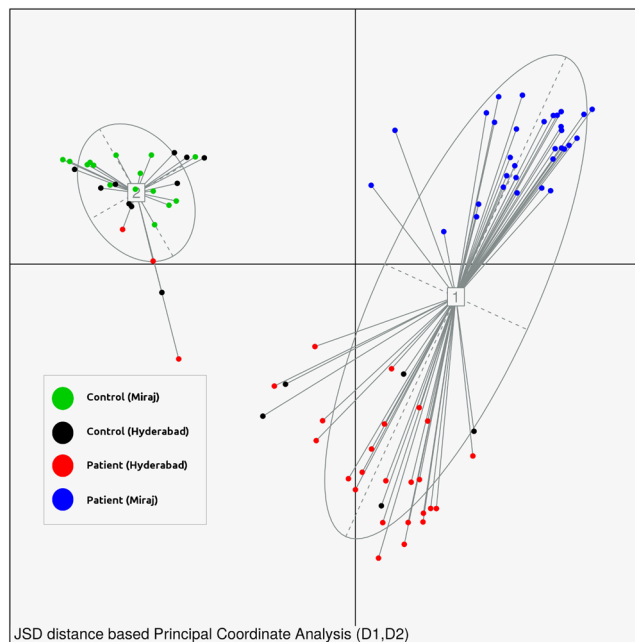


Fig. 5 JSD based PCoA (Ordination) plot for all samples of the study, colored by the nature (control or patient) and geographical affiliation (Miraj or Hyderabad) of samples.

from leprosy patients and healthy individuals using Illumina MiSeq sequencing. We have shared fastq data to enable re-analysis by end-users who intend/ prefer to employ customized computational protocols and/ or parameters of their choice.

Study metadata. In addition to microbiome data corresponding to sampled skin sites, the study additionally recorded Age, Sex, Disease Spectrum and Diagnosis record as aspects of the participating subjects. This is provided as a tab-separated file titled, 'Metadata for Leprosy Microbiome Study' that would facilitate customized data analysis in future retrospective studies.

Taxonomic abundance tables. As mentioned previously, we provide RDP classified taxonomic abundance profiles⁵⁴ at all five levels of taxonomic lineage and two distinct de-novo classification based taxonomic abundance tables. RDP classified datasets may be accessed through the data files titled:

1. Leprosy Microbiome **Genus** Level Abundance Data
2. Leprosy Microbiome **Family** Level Abundance Data
3. Leprosy Microbiome **Order** Level Abundance Data
4. Leprosy Microbiome **Class** Level Abundance Data
5. Leprosy Microbiome **Phylum** Level Abundance Data

Data files corresponding to the de-novo classified OTUs may be accessed through the files titled:

1. Leprosy Microbiome OTU Level Abundance Data (**Unfiltered, raw set**)
2. Leprosy Microbiome OTU Level Abundance Data (**Filtered Set**)

Data file corresponding to the image representing (Swab) sample collection process may be accessed through the file titled:

1. Sample Collection Image

Data file corresponding to the figure representing a dendrogram embedded heatmap based on genus level taxonomic abundances in all samples of the study may be accessed through the file titled:

1. Dendrogram Embedded Heatmap

Technical Validation

Ethical clearances were obtained from IEC, NCCS and leprosy clinics for performing experiments and implementing quality testing for the study. All the informed consents and questionnaire for counseling were precisely documented. The required questionnaire was bilingual in nature. Each individual enrolled in the study was rehearsed with the complete study design and protocols.

Sample management. Data entry was done in compliance with the data records obtained from the leprosy clinics. Multiple checks were done regularly to ensure its authenticity and validity.

DNA sample and PCR amplification quality control. The amount of DNA extracted from skin swabs was in extremely low amounts. Therefore, an aliquot of DNA extracted from each swab was used for PCR amplification using Touch-down protocol. PCR was performed in a Laminar flow cabinet after overnight UV illumination. The amplified product was then quantified using NanoDrop® ND-1000 UV-Vis Spectrophotometer (Thermo Fischer Scientific, USA) and checked for its quality using agarose gel electrophoresis. The PCR product was further outsourced for the amplicon library preparation and sequencing.

Sequence quality assessment and bioinformatics pipeline. Out of the total 85,57,782 good quality reads employed for reference based taxonomic classification, only 6,08,313 (~7%) remained unclassified. A fairly good percentage (55,38,880 i.e. ~65%) of the good quality reads were assigned to the most specific (genus) level using RDP classifier. In addition, a total of 161,695 OTUs were identified using vsearch based de-novo classification, which were filtered to 28,321 unique OTUs by omitting the population-level singletons.

Received: 28 May 2019; Accepted: 23 August 2019;

Published online: 22 October 2019

References

- Dethlefsen, L., McFall-Ngai, M. & Relman, D. A. An ecological and evolutionary perspective on human-microbe mutualism and disease. *Nature*. **449**, 811–818 (2007).
- Fredricks, D. N. Microbial ecology of human skin in health and disease. *J. Investig. Dermatol. Symp. Proc.* **6**, 167–169 (2001).
- Lunjani, N., Hlela, C. & O'Mahony, L. Microbiome and skin biology. *Curr. Opin. Allergy Clin. Immunol.* **19**, 328 (2019).
- Holland, K. T. & Bojar, R. A. Cosmetics: what is their influence on the skin microflora? *Am. J. Clin. Dermatol.* **3**, 445–449 (2002).
- Niemeyer-van der Kolk, T., van der Wall, H. E. C., Balmforth, C., Van Doorn, M. B. A. & Rissmann, R. A systematic literature review of the human skin microbiome as biomarker for dermatological drug development. *Br J Clin Pharmacol* **84**, 2178–2193 (2018).
- Leyden, J. J., McGinley, K. J., Nordstrom, K. M. & Webster, G. F. Skin microflora. *J. Invest. Dermatol.* **88**, 65s–72s (1987).
- Sanschagrin, S. & Yergeau, E. Next-generation sequencing of 16S ribosomal RNA gene amplicons. *J. Vis. Exp. JoVE*, <https://doi.org/10.3791/51709> (2014).
- Watts, G. S. *et al.* 16S rRNA gene sequencing on a benchtop sequencer: accuracy for identification of clinically important bacteria. *J. Appl. Microbiol.* **123**, 1584–1596 (2017).
- Klymiuk, I., Bambach, I., Patra, V., Trajanoski, S. & Wolf, P. 16S Based Microbiome Analysis from Healthy Subjects' Skin Swabs Stored for Different Storage Periods Reveal Phylum to Genus Level Changes. *Front. Microbiol.* **7**, 2012 (2016).
- Chakravorty, S., Helb, D., Burday, M., Connell, N. & Alland, D. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *J. Microbiol. Methods* **69**, 330–339 (2007).
- Grice, E. A. & Segre, J. A. The skin microbiome. *Nat. Rev. Microbiol.* **9**, 244–253 (2011).
- Grice, E. A. *et al.* A diversity profile of the human skin microbiota. *Genome Res.* **18**, 1043–1050 (2008).
- Zeeuwen, P. L. J. M. *et al.* Microbiome dynamics of human epidermis following skin barrier disruption. *Genome Biol.* **13**, R101 (2012).
- Belkaid, Y. & Segre, J. A. Dialogue between skin microbiota and immunity. *Science*. **346**, 954–959 (2014).
- Zeeuwen, P. L. J. M., Kleerebezem, M., Timmerman, H. M. & Schalkwijk, J. Microbiome and skin diseases. *Curr. Opin. Allergy Clin. Immunol.* **13**, 514–520 (2013).
- Cianci, R., Pagliari, D., Piccirillo, C. A., Fritz, J. H. & Gambassi, G. The Microbiota and Immune System Crosstalk in Health and Disease. *Mediators Inflamm.* **2018**, 2912539 (2018).
- Weyrich, L. S., Dixit, S., Farrer, A. G., Cooper, A. J. & Cooper, A. J. The skin microbiome: Associations between altered microbial communities and disease. *Australas. J. Dermatol.* **56**, 268–274 (2015).
- Balato, A. *et al.* Human Microbiome: Composition and Role in Inflammatory Skin Diseases. *Arch. Immunol. Ther. Exp. (Warsz.)*. **67**, 1–18 (2019).
- Silva, P. E. *et al.* Leprous lesion presents enrichment of opportunistic pathogenic bacteria. *SpringerPlus*. **4**, 187 (2015).
- Lastória, J. C. & Abreu, M. A. M. Mde Leprosy: review of the epidemiological, clinical, and etiopathogenic aspects - Part 1. *An. Bras. Dermatol.* **89**, 205 (2014).
- Eidt, L. M. Breve história da hanseníase: sua expansão do mundo para as Américas, o Brasil e o Rio Grande do Sul e sua trajetória na saúde pública brasileira. *Saúde E Soc.* **13**, 76–88 (2004).
- Han, X. Y. *et al.* Comparative Sequence Analysis of *Mycobacterium leprae* and the New Leprosy-Causing *Mycobacterium lepromatosis*. *J. Bacteriol.* **191**, 6067 (2009).
- Han, X. Y., Sizer, K. C., Velarde-Félix, J. S., Frias-Castro, L. O. & Vargas-Ocampo, F. The Leprosy Agents *Mycobacterium lepromatosis* and *Mycobacterium leprae* in Mexico. *Int. J. Dermatol.* **51**, 952 (2012).
- Franco-Paredes, C. & Rodriguez-Morales, A. J. Unsolved matters in leprosy: a descriptive review and call for further research. *Ann. Clin. Microbiol. Antimicrob.* **15**, 33 (2016).
- Jacob, J. T. & Franco-Paredes, C. The Stigmatization of Leprosy in India and Its Impact on Future Approaches to Elimination and Control. *PLoS Negl. Trop. Dis.* **2**, e113 (2008).
- Ridley, D. S. & Jopling, W. H. Classification of leprosy according to immunity. A five-group system. *Int. J. Lepr. Mycobact. Dis. Off. Organ Int. Lepr. Assoc.* **34**, 255–273 (1966).
- O'Neill, A. M. & Gallo, R. L. Host-microbiome interactions and recent progress into understanding the biology of acne vulgaris. *Microbiome*. **6**, 177 (2018).
- Xu, H. & Li, H. Acne, the Skin Microbiome, and Antibiotic Treatment. *Am. J. Clin. Dermatol.*, <https://doi.org/10.1007/s40257-018-00417-3> (2019).
- Ganju, P. *et al.* Microbial community profiling shows dysbiosis in the lesional skin of Vitiligo subjects. *Sci. Rep.* **6**, 18761 (2016).
- Nakatsuji, T. *et al.* Antimicrobials from human skin commensal bacteria protect against *Staphylococcus aureus* and are deficient in atopic dermatitis. *Sci. Transl. Med.* **9**, eaah4680 (2017).
- Paller, A. S. *et al.* The microbiome in patients with atopic dermatitis. *J. Allergy Clin. Immunol.* **143**, 26–35 (2019).
- Alekseyenko, A. V. *et al.* Community differentiation of the cutaneous microbiota in psoriasis. *Microbiome* **1**, 31 (2013).
- Chiu, L. *et al.* Protective Microbiota: From Localized to Long-Reaching Co-Immunity. *Front. Immunol.* **8**, 1678 (2017).
- Silva, P. E. S. *et al.* Insights into the skin microbiome dynamics of leprosy patients during multi-drug therapy and in healthy individuals from Brazil. *Sci. Rep.* **8**, 8783 (2018).
- Balcells, M. E., Yokobori, N., Hong, B., Corbett, J. & Cervantes, J. The lung microbiome, vitamin D, and the tuberculous granuloma: A balance triangle. *Microb. Pathog.* **131**, 158–163 (2019).

36. Rao, P. N. & Suneetha, S. Current Situation of Leprosy in India and its Future Implications. *Indian Dermatol. Online J* **9**, 83–89 (2018).
37. Moreno, J. D., Schmidt, U. & Joffe, S. The Nuremberg Code 70 Years Later. *JAMA*. **318**, 795–796 (2017).
38. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*. **310**, 2191–2194 (2013).
39. Kaiser, K. Protecting Respondent Confidentiality in Qualitative Research. *Qual. Health Res.* **19**, 1632–1641 (2009).
40. Rhodes, R. Ethical issues in microbiome research and medicine. *BMC Med.* **14**, 156 (2016).
41. Kong, H. H. Details Matter: Designing Skin Microbiome Studies. *J. Invest. Dermatol.* **136**, 900–902 (2016).
42. Kong, H. H. *et al.* Performing Skin Microbiome Research: A Method to the Madness. *J. Invest. Dermatol.* **137**, 561–568 (2017).
43. Schmieder, R. & Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics*. **27**, 863–864 (2011).
44. Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. PEAR: a fast and accurate Illumina Paired-End reAdmergeR. *Bioinformatics*. **30**, 614–620 (2014).
45. Cole, J. R. *et al.* The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res.* **37**, D141–145 (2009).
46. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ*. **4**, e2584 (2016).
47. Hsieh, T. C., Ma, K. H. & Chao, A. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods Ecol. Evol.* **7**, 1451–1456 (2016).
48. Ito, K. & Murphy, D. Application of ggplot2 to Pharmacometric Graphics. *CPT Pharmacomet. Syst. Pharmacol.* **2**, e79 (2013).
49. Dhariwal, A. *et al.* MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucleic Acids Res* **45**, W180–W188 (2017).
50. Arumugam, M. *et al.* Enterotypes of the human gut microbiome. *Nature* **473**, 174–180 (2011).
51. Langille, M. G. I. *et al.* Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* **31**, 814–821 (2013).
52. Nagpal, S., Haque, M. M., Singh, R. & Mande, S. S. iVikodak—A Platform and Standard Workflow for Inferring, Analyzing, Comparing, and Visualizing the Functional Potential of Microbial Communities. *Front. Microbiol.* **9**, 3336 (2019).
53. *NCBI Sequence Read Archive*, <http://identifiers.org/ncbi/insdc.sra:SRP187334> (2018).
54. Bayal, N. *et al.* 16S rDNA based skin microbiome data of healthy individuals and leprosy patients from India. *figshare*, <https://doi.org/10.6084/m9.figshare.c.4478453> (2019).

Acknowledgements

This study was funded by Department of Biotechnology, Government of India under the SysTB project (DBT/PR3260/BRB/10/967/2011) and the Centre of Excellence grant (DBT/PR15450/COE/34/46/2016). Authors would like to thank the administration and staff of Lepira Blue Peter Public Health and Research Centre (Hyderabad) as well as Richardson Leprosy Hospital, The Leprosy Mission Trust (Miraj) for their support. The subset of authors from Tata Consultancy Services Limited (TCS) who were involved in aspects of study design, analysis of data, interpretation of results, and drafting of the final manuscript thank their organization (TCS) for providing funding support in form of monthly remuneration (salaries) and required computational infrastructure for performing the analysis. Nitin Bayal is a PhD fellow at NCCS under the DBT-JRF fellowship - Department of Biotechnology, Government of India, and would like to acknowledge NCCS and DBT for its support.

Author contributions

S.C.M. conceptualized the study. N.B., M.S.P. and S.C.M. designed the sampling protocols and time-lines. V.V. and R.S. assisted development of sampling strategy and management of records. N.B. collected and processed samples. S.N. and N.B. performed computational analysis. S.C.M., S.S.M., M.M.H. supervised computational work. S.N. designed figures and graphs. M.M.H., S.N., N.B., M.S.P. wrote manuscript. S.C.M. and S.S.M. supervised and proofread the manuscript. All authors read and approved the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.C.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

The Creative Commons Public Domain Dedication waiver <http://creativecommons.org/publicdomain/zero/1.0/> applies to the metadata files associated with this article.

© The Author(s) 2019