Predominance of Extreme Environmental Conditions at High Altitude in Modulating Human Gut Microbiota

Brij Bhushan[1], Malleswara Eslavath[1], Anand Yadav[1], Ashish Srivastava[1], Maran Prasanna Reddy[1], Tsering Norboo[2], Bhuvnesh Kumar[1], Lilly Ganju[1], and Shashibala Singh[3]

[1]DRDO Defence Institute of Physiology and Allied Sciences
[2]Ladakh Institute for Prevention (LIP)
[3]National Institute of Pharmaceuticals Education and Research (NIPER)

August 3, 2020

Abstract

Human microbial alterations are associated with environmental stress, nutritional, genetic and triggering de-novo variations. Nevertheless, human gut microbiome at extreme altitude (>5800 m) remains unexplored. We aimed to demonstrate the microbial predominance in individuals with same ethnicity and dietary pattern at extreme altitude with unique challenges like cold, hypoxia, radiation etc. Different analysis pipelines were used for fecal whole genome sequencing at 210m, 3500m, 4420m and 5805m, and 16s rRNA V3-V4 regions amplification sequencing of 19 individuals belonging to the same ethnicity and dietary pattern, for presence of taxonomy & functional potential and confirming the prediction upto the strain level within the same cohort. Principal component analysis, revealed distinct microbiome changes at different altitudes, with varied and higher Bacteroides and Prevotella ratio. There was predominance of genus Prevotella at altitudes 4420m & 5805m than at 210m & 3500m. Appearance of species Prevotella copri strain 61740 was increasing significantly at extreme altitudes, whereas co-occurrence of other bacterial strains had different pattern than Prevotella. The extensive strain level analysis indicated alteration in the metabolic pathways. This study under stressful and hypoxic environment of extreme altitudes, associated microbial variation with altered metabolic pathways, reveals influence of extreme environment on human gut microbiota with predominance of Prevotella.

Introduction

The high-altitude (HA) area in the Karakoram range has typical environmental challenges, such as low oxygen pressure, temperature, humidity, high radiation and dry weather etc., leading to series of physiological and metabolic disorders[1,2] . In addition, change in the diet includes high fiber and low protein products as major portion of diet, while consumption of vegetables and fruits is significantly reduced. Such places are mainly populated with natives, personnel on duties and researchers. Several associations between chronic human diseases including inflammatory bowel disease (IBD) like Crohn's disease, some extra-intestinal disorders including asthma, metabolic syndrome, allergies, cardiovascular disease, and obesity [3,4,5,] , altered gut microbiota constitution and its functions have been reported in last few years[6,7] . At least 30% of sojourners from low land report adverse effects at HA including cardiac, respiratory & neurological symptoms, joint pains, headache, nausea, vomiting, bloating and IBD. Little work has been done on the effect of HA hypoxia on digestive system in either patients with altitude illness or in healthy individuals. In fact, symptoms of digestive system such as anorexia, loss of appetite, stomach/abdominal pain, constipation, epigastralgia, heart burn, dyspepsia, vomiting, diarrhea, hematemesis, piles and peptic ulcers are frequently reported in mountaineers and altitude sojourners[8,9,10,11,12] . Moreover, epidemiological and clinical studies suggest that gastro Intestinal bleeding (GIB) is not uncommon at HA, and is often life-threatening[13,14] . The problems

could be due to the release of host metabolites causing inflammation or the interaction between the host and gut microflora[14,15,16] .

Exciting time to study the body associated microbes and the human microbiota, consisting of 100 trillion bacteria along its 400m^2 surface area, encoding 100fold more genes than our own genome[17] . The in-depth analysis of the microbes make it to realize their influence on human physiology in different areas to characterize the global presence of microbial diversity under different environments and dietary habits. It is reported that the microbial diversity in the gastrointestinal tract plays a crucial role in host physiology, including involvement in nutrient metabolism and gut barrier[18] . However microbiome composition driving factors are not fully understood yet, due to several factors including limited sample size, inconsistent DNA extraction and varied sequencing methods[19,20,21] . Various reports support the involvement of environment, genetics, antibiotic usage[22] and lifestyle factors including diet[23,24,25] . Even short exposure to HA can alter the composition of gut microbiota[26,27] . Li and Zhao[28] and Li et al[29] , have reported contribution of HA, genetics and dietary factors in Tibetan and Chinese Han groups in shaping gut microbiota. Overall gut microbiome has been investigated in some Indian tribal populations irrespective of their ethnicity and dietary habits[20] . Das et al[30] compared Indian urban and rural population at 228m with HA population at 3500m, with respect to their dietary habits and geographies. Study performed on western Indian population at 153m compared to other geographies revealed a distinct taxonomic and functional niche[31] .

To the best of our knowledge there has been no study so far on genome wide changes on human response at extreme altitude and that ours is the first attempt as a comparative analysis on Indian sojourners belonging to the same ethnicity ascending to extreme altitude, to compass the complete meta genomic map and networks involved in extreme altitude (3500 – 5800m) responses. We used 16s rRNA associations with Whole genome sequencing (WGS) applying the same metagenomics DNA extraction & sequencing platform and bioinformatics pipeline to yield not only profiles of bacterial composition and diversity, but also estimate the functional potential of the microbiome[32] . For quantifying bacterial species abundance and intra-species genomic variation we used MIDAS, a computational pipeline, followed by novel strain level analysis by METABET2 approach to quantify gene content of prevalent bacterial species and identified significant intra species population composition associated with environmental factors at extreme altitudes. The findings revealed striking microbial dysbiosis proving important to shed light on human acclimated processes and prevention of the physiological changes associated with it at HA. It is expected that besides classifying the gut microbial repertoire, particularly going down to strain variation at nucleotide level and differences in gene content, may provide crucial information in deciphering important insights in correlating the structure of gut microbiome, microbial pathogenicity, various metabolic disorders and host adaptation of pathogens[33] or other related microbial conversions and outbreaks[34,35] at high and extreme altitudes.

Materials and Methods

Subject recruitment, enrolment and processing

Nineteen subjects were recruited from a group of HA sojourners, from North Indian Ethnicity who had not travelled above 3000m in past six months. Inclusion criteria were fasting, no antibiotic treatment, no known gastro-intestinal diseases and previous bariatric surgery or medication known to affect the immune system and metabolism. All the participants were males in the age group of 22-55 years, had normal weight (BMI = 20-24kg /m2). Everyone was subjected to undergo thorough medical and psychological examinations for any diseases to ensure the healthy population. Information on medication status was obtained by questionnaire and interview on the first day of examination.

The analysis of gut microbiota of 19 sojourners from North Indian Ethnicity, with similar dietary habits, who climbed from 210m base line (H1) to various heights of Korakuram range, was performed using Illumina Miseq platform. Three participants, after staying at 4420m (H2) for six months descended to 3500m (H3). After staying for two months now ascended to 5800m (H4) to stay for four months. Thus 12 fecal samples from three subjects at four heights were analyzed for WGS. Whereas 32 samples of remaining 16 participants from the same cohort were analyzed by 16S rRNA sequencing for two heights (H1 and H2), to illustrate

microbial diversity at extreme altitudes.

Participants collected stool samples in 1 ml of RNA later early in the morning. The samples were immediately distributed in aliquots of 500 µl each, and stored at -80°C for sequencing analysis. Though there is evidence that storage of samples at low temperatures maintains the microbial community structure[36] , but yet the time span from sampling to delivery was intended to be as short as possible[37] .

Statement and volunteer information

All the participants were well informed about the study protocol, the sampling procedure and research method. The written consent was obtained after they understood the nature and relevance of the study. The ethics committee of Defence Institute of Physiology and Allied Sciences, Defence Research and Development Organization, Delhi, India, granted the study approval. The study protocols were in accordance with the approved Helsinki guidelines.

DNA extraction and whole-genome shotgun sequencing

Microbial DNA extraction from the fecal samples was performed using the Power Fecal DNA Isolation Kit (Mobio, USA), according to manufacturer's protocol. WGS library was constructed according to standard True Seq WGS protocol provided by Illumina, Inc. (San Diego, USA). Library quality for fragment size distribution was checked on Bioanalyzer 2100 using high sensitivity DNA kit (Agilent technologies, USA). Library quantification was performed using a Qubit Fluorometer (Invitrogen, USA) and a Stratagene Mx3000P Real-time PCR Cycler (Agilent, USA) prior to cluster generation in a c-Bot automated sequencing system (Illumina, Inc.).

16S rRNA amplicon Library Preparation

Illumina MiSeq platform was used to perform 16S rRNA amplicon sequencing using the protocol described by Caporaso et al in 2010[79] . Briefly, we used universal primers 341F and 805R to amplify V3-V4 hypervariable region of 16S rRNA gene. The target amplified product size was approximately 464 bp. The V3V4 primer and the adapter details are as mentioned below.

V3V4 amplification primers

Primer ID	Primer Sequence
341F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG
805R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC

The final amplified amplicon libraries were purified using AMPure XP beads (Beckman Coulter Genomics, Denver, MA, USA) and the size and quantity of the amplicon library were assessed on the bioanalyser (Agilent technologies, USA) & the library quantification kit for illumina (Kapa Biosciences, Woburn, MA, USA), respectively.

Sequencing WGS and 16S amplicon libraries

Twelve WGS DNA libraries with different indices were diluted & pooled together with an equimolar concentration of each library. The pooled library was denatured and loaded on HiSEQ 2500 flow cell lanes for cluster generation in a c-Bot automated sequencing system (Illumina, Inc.), using illumina Cluster Generation kits. The sequencing was performed using TruSeq SBS v3 reagent kit (200cycles) for 100 base paired runs.

The 16S libraries were sequenced on MiSEQ platform (Illumina CA, USA). PhiX Control library was combined with the pooled V3-V4 amplicon library (expected at 20%) and loaded on to MiSEQ flow cell for clustering, to a density of approximately 570K/mm^2. For sequencing MiSEQ reagent kit (500cycle) was used to generate 250 bp PE reads.

16S rRNA Metagenome analysis

Raw paired-end Illumina reads were trimmed for 16S rRNA gene primer sequences using Cutadapt (https://cutadapt.readthedocs.org/en/stable/). Pre-processed sequence data were filtered using the Quantitative Insights Into Microbial Ecology (QIIME), software suite. Sequences were clustered into OTUs on the basis of similarity to known bacterial sequences available in the Greengenes database (v13.8; http://greengenes.secondgenome.com/;) (at 97% sequence similarity cut-off) using UCLUST clustering software. Sequences that could not be matched to references in the Greengenes databases were clustered de novo based on pair-wise sequence identity. Total sum scaling (TSS) normalization was applied followed by CSS to correct biases introduced by TSS, and log2 transformation to account for the non-normal distribution of taxonomic counts. Statistical analyses were executed using the Calypso software V8.72[41](cgenome.net/calypso/).

Statistical Analysis 16S rRNA and WGS

Samples were clustered using supervised canonical correspondence analysis (CCA) (including 'Variable heights' H1 to H4 as explanatory variables). The community richness was identified by rarefraction analysis for each sample using Calypso pipeline (Zakrzewski et al 2016). Differences in bacterial alpha diversity (Shannon diversity, Richness, Evenness, Chao & Fischer Alpha) between groups were calculated.

Differentially abundant and significant taxa were evaluated using Wilcoxon rank test (odds ratio) at a p-value cut off of 0.05 followed by false discovery rate (FDR) correction of the p value to determine statistical significance (i.e., $p < 0.05$/number of tests) for reproducibility. We also generated quantile – quantile plots (QQ plots) of the—log (observed p value) versus the—log (expected p values) within each pair wise comparison for all taxonomic levels and gene categories to ascertain potentially statistically significant associations after correction for multiple comparisons.

Beta diversity was calculated using weighted UniFrac distances, whereas, differences were calculated using Permutational Analysis of Multivariate Dispersions (PERMDISP2) through the betadisper function. Differences in the composition of the fecal microbiota between groups were assessed using the Linear discriminant analysis Effect Size (LEfSe) workflow, by comparing each height (H2, H3 & H4) to baseline (H1).

Whole genome sequencing Metagenome analysis

The WGS Metagenome data was processed using Metagenomic analysis tool kit, Version 2 (MOCAT2),[38](http://vm-lux.embl.de/~kultima/MOCAT/). The raw reads were filtered by removing low quality reads, followed by mapping of high quality filtered reads to the*RefMG.v1 database (* http://www.bork.embl.de/software/mOTU/) and the*mOTU.v1 database* (http://www.bork.embl.de/software/mOTU/download.html)[39] . These databases were used to generate taxonomic profiles using single copy marker gene method or metagenomic operational taxonomic units-mOTU, from phylum to species for all samples. Data transformation was performed using Cumulative-sum scaling (CSS), a widely used method for normalizing microbial community composition data[40] .The data was further transformed to log2 to account for the non-normal distribution of taxonomic counts. Statistical analyses were executed using the Calypso software V8.72[41] (cgenome.net/calypso/).

Metagenome species and the strain level analysis

As it is known that the human health conditions are linked to microbial communities, phenotypes are often associated with only a subset of strains within causal microbial groups. Therefore, for WGS metagenome data we used Metagenome binning with abundance and tetranucleotide frequencies V.2 (Metabat2)[42] , and Metagenomic Intra-Species Diversity Analysis System (MIDAS)[43] , tools for identifying metagenome species and strain-level metagenomic classification at default parameters. De novo assembly for all 12 samples was performed by using short oligonucleotide analysis package (SOAP)[44] , at K-mer65, followed by binning using Metabat2 software. Bins greater than 150 genes were selected for further analysis. Genes which were differentially abundant and had $p<0.05$, were considered for visualization.

Species-level coverage was obtained by using MIDAS database across samples from sojourners visiting dif-

ferent heights. For species with sufficient coverage, reads were aligned to a pan-genome database of genes to estimate gene coverage[45] , copy number and presence or absence. The core genome was defined directly from the data by identifying high-coverage regions (>70% coverage of the pangenome genes), across multiple metagenomic samples, providing a comprehensive strain-level genetic overview of the gut microbial diversity.

Functional Metagenome & Pathway Analysis

For functional analysis, MOCAT2 was used to cluster high quality reads into reference gene catalogs which is then annotated to Evolutionary genealogy of genes: non-supervised orthologus group (eggnog) database[46] . The gene abundance profile was calculated for each sample by mapping the raw reads back to the reference gene catalogue. Functional mapping and Analysis Pipeline (FMAP)[47] , was used to compare the various time points for alteration in pathways, using the Kruskal-Wallis rank-sum test.

Results

Nineteen participants climbed from 210m (H1) to 4420m (H2) to stay for six months. Three of them descended to 3500m (H3), stayed for two months to further ascend to 5800m (H4) for four months. Fecal samples collected at different time points were analyzed for WGS (for three subjects who climbed to H4 and compared with 16S rRNA sequencing of subjects (16) from the same cohort who stayed at H1 and H2 heights to illustrate microbial diversity upto strain level and functional analysis at extreme altitudes.

Sharing of Sequence data between four different altitudes

Raw data statistics

The total sequence data of 12 stool samples from three sojourners at four time points were 20 Million average reads and an average read length 2x150 bp (Supplementary Figure S1(a)). Stool samples from remaining 16 participants (32 samples) were analyzed by 16S rRNA sequencing for H1 and H2 heights (Supplementary Figure S1 (b)). Using MiSEQ an average of reads and read length 126000 and 2x250 bases, respectively, was generated.

The raw abundance profile from WGS data (MOTUs) of 12 samples (Supplementary Figure S2(a) and abundance profile of OTUs from 16s rRNA data of 32 samples were normalized and transformed by using TSS (CSS+Log2) (Supplementary FigureS2(b)).

Statistical Analysis 16S & WGS datasets: Alpha & Beta diversity (PcoA), Differential Abundance & PERMANOVA

Various bacterial diversity metrics were employed for inferring the structural aspects of the microbial community. The alpha diversity was calculated to determine by Shannon index & Fisher's alpha, Chao richness and evenness (Figure 1A). The results of both16s amplicon and shot gun metagenome demonstrated similar diversity patterns at height H2 (Figure 1B). Other comparisons indicate major diversity shifts at all heights. The rarefraction curve of all samples calculated had reached a plateau, suggesting that the sequencing was deep enough at phylum and genus levels (Figure 1C). Good's coverage estimation revealed 80.3% to 91.0% species in all the samples.

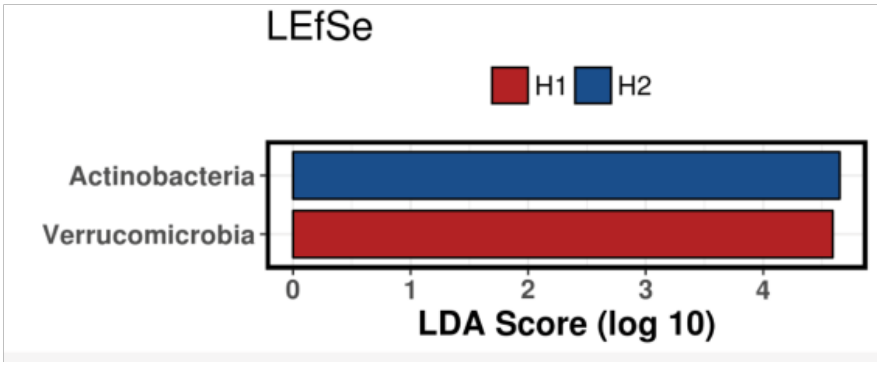
Supervised PCA and CCA results (Figure 2), demonstrate separate clusters (significant differential distribution of fecal microbiota at $p = 0.05$), between altitudes H1 & H2 employing both 16s rRNA (Figure 2 Ai and Aii) and WGS methods (Figure 2 Bi, Bii). Global microbial community composition using Permutational Analysis of Multivariate Dispersions (PERMDISP) 2 clustering, also indicated significant differential distribution between H1 and H2 heights (Figure2Aiii&Biii), as well as all four heights H1, H2, H3 and H4 (Figure 2 C & D).

Comparative taxonomy-based gut microbiota at the phylum and genus levels at different altitudes

Phylogenetic classification analysis of the metagenomic reads for each sample was performed using the Greengene 13.5 database. The results were assigned at phylum and genus levels based on an identity level

of 97% similarity.

The taxonomic analysis at Phyla level



The structure of sojourners gut microbiota was analyzed at different heights, measuring the abundance of bacteria by WGS, representing known sequences at phyla level more than 1% using Wilcoxon rank test (odds ratio at 95% confidence intervals). There were eight phyla present in the fecal samples at different heights (Figure 3). Phyla Bacteriodetes was predominantly present, accounting for 80% of the identified taxa at H1 time point. The second richest phyla was Fermicutes with 18% presence. The next detectable phyla was 2% Proteobacteria (Figure 3 A). After ascending to and six months stay at H2, there was a significant decrease in Bacteriodetes to 64%, though still remaining the richest phyla, increase in Fermicutes contributing 35% with Proteobacteria remaining the same. The microbial composition varied immensely, depicting an interesting phenomenon when sojourners descended to H3 and stayed for two months. There was a sudden 5% appearance of Fusobacteria and increase in Proteobacteria to 6%. Moreover, the predominance of Bacteriodetes was reversed and increased to 76%, whereas Fermicutes reduced to 8%. At the end when sojourners ascended to H4 and stayed for four months, the newly appeared phyla Fusobacteria at H3, suddenly disappeared remaining with Bacteriodetes, Fermicutes and Proteobacteria 77%, 13% and 3% respectively. Other bacterial members from Actinobacteria, Verrucomicrobia, Lentisphaereae, Euryarchaeota were also detected with lower abundance.

Further, confirmation of the variation in the composition of fecal microbiota of sojourners at different heights was achieved by linear discriminant analyses (LEfSe) using the Greengene software (http://greengenes.lbl.gov). FDR correction of the p value was to determine statistical significance ($p < 0.05$/number of tests) and considered $p < 0.05$ statistically significant for reproducibility. Figure 3D indicates the association between the relative abundance of presence or absence of phyla and the significant change in the microbiota mainly with abundance of Bacteriodetes ($p= 0.05$), Fermicutes ($p= 0.05$) and Proteobacteria ($p=0.05$). LEfse analysis also showed concordant result with Wilcoxon rank test (Supplementary FigureS3). The results thus indicate the most important contributing factor to the diversity could be altitude or hypoxia as diet and ethnicity of all the subjects were the same.

The presence of eight phyla by WGS analysis in limited three subjects were reproduced by 16s RNA analysis in another 16 subjects from the same cohort at H1 and H2 heights ($p<0.05$ Wilcoxon rank test) with odds ratio of 95% (Figure 3B). The second method reconfirmed the significant changes in the abundance of phyla Bacteriodetes and Fermicutes (Figure3C).

The taxonomic analysis by WGS at Genus level

Next, we analysed the frequency of most abundant genera in 12 fecal samples at four different altitudes (Figure 4). The relative abundance of around 20 genera was more than 1%. At H1 time point genus Bacteroides was the most abundant (46%), followed by Prevotella (19%), Parabacteroides and Alistipes (7%), Eubacterium (5%), Faecalibacterium (3%) Ruminococcus (2%), Blautia, Coprobacillus, Clostridium, Escherichia,

Bifidobacterium, and unclassified (1%) (Figure 4 A). Interestingly, after ascending to 4420m (H2) and staying for 6 months, the richest genus was Prevotella (63%) with significant decrease in Bacteroides to 2%, followed by Eubacterium, Faecalibacterium 3%, and Parabacteroides 1%. There was a marginal increase in Coprococcus contribution to 4%. Surprisingly even after getting adapted to stressful environment at H2, on descending to H3, the scenario changed markedly where Bacteroides reversed and increased significantly to 31%, though still lower than Prevotella which was reduced to 43%. Incidently, there was sudden appearance of Fusobacterium (5%), Roseburia (4%) and Escherichia (2%). The pattern was similar to H1. Interestingly, again the phylogeny structure was changed when they ascended to 5800m (H4) and stayed for four months. Bacteroides significantly reduced to 9%, Prevotella significantly increased to 65%, Eubacterium (3%), Faecalibacterium (2%), whereas Parabacteroides, Alistipes, Coprococcus, Escherichia, and Coprobacillus (1%) started reappearing. The pattern looked similar to H2. When analyzing the overlap of different genera at four different altitudes, it mostly belonged to Bacteroides and Prevotella, where Prevotella covered the 63% of the total genera at H2 and H4 heights and Bacteroides at H1 and H3. Although 3500m (H3) is defined as high altitude, interestingly, the pattern of the metagenomic profile obtained at H3 was quite similar to the base line profile (H1= 210m).

Further confirmation of the differences in the composition of fecal microbiota of sojourners at different heights was achieved by LEfse using the Greengene software. FDR correction of the p value was to determine statistical significance ($p < 0.05$/number of tests) and considered $p < 0.05$ to be statistically significant for reproducibility (Figure 4D). Thus the genus abundance clearly indicated the significant change in the microbiota mainly with abundance of Bacteroides ($p= 0.05$) and Prevotella ($p= 0.05$). LEfse analysis also showed concordant result with Wilcoxon rank test (Supplementary FigureS4 with LDA score) indicating altitude or hypoxia, could be the most important contributing factor to the diversity.

The WGS analysis of three subjects revealed the presence of 20 genera. Amongst these, the relative abundance of genus Bacteroides, Prevotella, Bifidobacterium and Fecalibacteria were significantly altered (with $p <0.05$ Kruskal Wallis H test) (Figure 4 A). We reproduced and correlated the significant presence of 20 genera in another 16 subjects of the same cohort at H1 and H2 heights ($p<0.05$ Wilcoxon rank test) with odds ratio of 95% (Figure 4B), by 16s rRNA analysis. The second method reconfirmed the changes in the abundance of genus Bacteroides, Prevotella, Faecalibacterium & Bifidobacterium as in WGS (Figure4 C). Interestingly compared to WGS, the 16s rRNA also had higher prevalence detection rate for other genus like Ruminococcus and Akkermansia at H2 height.

Co-abundance of Gene Groups and Metagenome Species/Strain with a consistent definition and efficient algorithm

De Filippis et al[48] identified that within the genera Prevotella, the presence of distinct oligotypes had differential associations with non-vegetarian and vegetarian diets[48] . In the present study, the representative sequence of OTUs assigned to Prevotella were mapped on to the known oligotypes to check the species association. About 90% of the Prevotella OTU sequences identified were P11 and P12 oligotype representative of non-vegetarian type, indicating their prevalence in North Indian population (Supplementary Table ST1)

To quantify species and strain level genomic variation accurately and broadly at four different heights, co-abundance of genes (CAG) were identified by binning of SOAP- a denovo assembly of all Shotgun Metagenome samples. A total 57 bins were obtained from Metabat2 software. 12 bins which had greater than 150 genes, were accurately selected for further processing to avoid erroneous, inconsistent and incomplete annotations that would affect some taxonomies. Genes which were differentially abundant and had $p <0.05$ were considered for visualization (Supplementary Figure S5).

We used an integrated pipeline for profiling both species and stain level abundance and genomic variations, from metagenomes. MIDAS analysis pipeline generated few more bacteria in addition to results obtained from Metabat2. MIDAS was able to capture the majority of microbial species abundance across the subjects, making it well suited for uncovering strain-level variation associated with various heights. For species with maximum coverage, reads were aligned to pan-genome database of genes to estimate gene coverage, copy

number and presence or absence and finally detected SNPs. The pangenome reconstructed bacterial profile revealed at all time points, was filtered for minimum pangenome coverage of 70 % demonstrating 37 species and strains (Supplementary Figure S6). The relative abundance profile generated was analysed through T-test to identify highly significant taxa with an FDR cut off of 0.05. There was a significant correlation between Metabat2 and MIDAS results which justifies the presence of Roseburia, Prevotella, Faecalibacterium, Eubacterium & Bacteroides, significantly enriched out of 37 species by both the methods of analysis.

Strain analysis

SNP level analysis was performed across the subjects by mapping metagenomic reads to bacterial reference genomes and quantify nucleotide variation along the entire genome (i.e. the read depth and observed alleles at each position), to identify multiple strains of bacteria playing role in adaptation. Further SNP nucleotide variants from all the samples were pooled and filtered at 15X depth at a minor allele frequency of 0.01.

Strain level maximum likely hood phylogenetic tree was constructed (FAST-TREE http://www.microbesonline.org/fasttree) (Figure5*),* using consensus alleles obtained from the core-genome sequences of Prevotella copri 61740 having nucleotide variation (multi-FASTA file in supplementary at H1, H3 to H2 and H4), which separates it from strain at H1, H3 versus H2 and H4 (High altitudes), respectively.

It is observed that the abundance of significantly appeared individual nine strains, Roseburia inulinivorans 61943, Faecalibacterium prausnitzii 62201, Oscillibacter sp 60799 and Eubacterium eligens 61678 had similar pattern of appearance as increase at H2, decrease at H3 and again increase at H4. Except Roseburia inulinivorans 61943 other three had a sharp decline at H3. Megamonas hypermegale 57114 and Eubacterium rectale 56927 decreased linearly at H2 and H3 with sudden increase at H4. Whereas, Bacteroides vulgatus 57955 showed a very different and declined pattern at H2, which persisted till H3, with a marginal increase at H4. Contrary to all these patterns Prevotella copri 61740 had sharp increment at H2 which further increased to H3 with a mild decrement at H4 (Figure 6).

Functional diversity analysis

In order to investigate the functional potential, we compared the abundance of specific pathways and functional processes across the samples at H2, H3 and H4 in comparison to H1, and inferred using MOCAT approach. FMAP software was used to identify differentially abundant (DA) genes using Kruskal-Wallis rank-sum test (fold change > 2, pvalue cutoff 0.05). Pathway enrichment analysis was used to determine differentially abundant pathways using Fishers exact test[49] , with average log2 (fold change) and enriched significance (p value cut-off of 0.01), to identify significant differentiators between the heights. The analysis identified a set of 27 different functional processes (Figure 7).

FMAP generated significantly noticeable differences between functional potential encoded at different heights, mainly down regulated Carbohydrate; Amino Acid; Energy; Vitamins & Cofactor, biosynthesis of secondary metabolites & antibiotics and other Metabolic Pathways at H2 and H4 as compared to H1 (Figure 7). Incidentally, these pathways were upregulated at H3, which was expected, as the sojourners descended from H2 to H3, a bit close to H1. Samples at H2 and H3 appeared to be more enriched with genera than at H4 that encoded functions associated with four other pathways namely Biosynthesis of Amino acids and Fatty acid metabolism, both at H2 and H3, whereas, Xenobiotic degradation & metabolism and endocrine and metabolic diseases are more at H2, and Lipid metabolism and Glycan biosynthesis & metabolism are over-represented at H3. Interesting to note that the microbial metabolism in diverse environments was markedly reduced at all three heights. The pathways detected in three subjects by WGS were reproduced by 16s rRNA in another 16 subjects at H1 and H2. The most common six sub-pathways were glucose seven phosphate isomerase; GDP mannose 4,6 dehydratase; 8-amino7-Oxononanoate Synthase; GMP synthase (glutamine hydrolysine); Phospho-ribosyl-formyl glycine amidinecycloligase; Preprotein translocase subunit fraction secA; Phospho-ribosyl-formyl-glycine–amidine-synthase (Supplementary Figure 8.) It is quite possible that these functional capabilities are related to high altitudes stress, however further studies are required.

Discussion

The present study was performed to analyze the gut microbiota of 19 Indian sojourners on an extreme altitude expedition. We compared two sequencing methods, WGS and 16srRNA to latch on the complete picture of the gut microbial diversity and functional networks involved at extreme altitudes under low oxygen pressure. By doing so, we uncovered both known and novel microbial components, upto the species and strain levels, including a highly altered network of various functions indicative of leading to pathological changes. The two analytical methods employing various state of art pipelines presently available for gut meta genomes, verified our meta-analysis framework to be robust and appropriate. To the best of our knowledge there has been so far no study on microbiome -wide expression changes in Indians high and extreme altitude responses and that ours is the first one and is therefore imperative in shedding light on some prospective cohort studies and pooling data at various altitudes with large sample size, that should be performed in order to study the effect on the human gut microbiome in identifying important associations that may be involved with detection and prevention of some pathological diseases.

The study represents significant microbial changes in individuals exposed to various altitudes which could be attributed to environment while with similar diet and genetic patterns. The encouraging part of the study was, samples were collected longitudinally making it possible to determine microbial changes associated with the induction to different altitudes and easily compared with their own baseline profile. However, the study was not without limitations. It had relatively small sample size, therefore, we were underpowered to detect many statistically significant details, which were nevertheless circumvented by using more than one methods and latest pipelines, providing additional information for down-stream data pooling and meta-analysis[50] . We were able to use 16S rRNA gene sequencing to reproduce significant bacterial abundance in the same population that was assessed using WGS. We chose to collect fecal samples mainly because of its easy accessibility, which is extremely important if changes in the microbiome revealed were to be used as indicators of the entire human health status.

Human gut is extraordinary dense with microbes, facing the greatest challenge of number & diversity of intestinal species[29,51] . There are many factors having impact on the composition of gut microbiota. However, so far, the statistical analysis indicated that the three most important contributing factors are environment, genetics and diet[29] . Hence, we chose to focus on examining the impact of altitude and related environmental conditions on the differences in the gut microbiota of Indian male sojourners of same age range, same ethnicity, altitude (210m H1) and similar dietary habits, visiting different altitudes (3500, 4420 to 5800m) for short time periods.

The gut microbiota of Indian sojourners was compositionally distinct at high and low altitudes. Three samples for WGS had similar diversity measures compared to another sixteen sequenced by 16s rRNA method. Overall Bacteriodetes and Firmicutes revealed their marked presence. A relatively lower abundance of Proteobacteria is apparently a positive state of gut health, indicating no epithelial dysfunction[52] in the recruited cohort. At sea level H1, the relative abundance of phyla Bacteriodetes was much more than Firmicutes, which is in line with the previous studies on Indians[19,20,21,30] .The same subjects after staying at H2 for six months had altered pattern with decreased Bacteriodetes, though still remaining the maximum. Whereas after descending to H3 and staying for two months, resulted in reversal of pattern by increase in Bacteriodetes & Proteobacteria, similar to the findings of Das et al at 3500m[30] and sudden appearance of Fusobacteria. Again, after ascending to H4 and stay for four months, Fusobacteria disappeared with increase in Fermicutes and decrease in Proteobacteria whereas Bacteriodetes remaining almost the same and maximum. All the study subjects had normal BMI as per the standards[53] and similar diet. Bacteriodetes have been reported to make distinctive lipids associated with atheroma formation and contribute to heart disease[54] . Therefore, to investigate the effect of different altitudes on gut microbiota of Indian subjects of same ethnicity with high abundance of Bacteriodetes, consuming similar diet and visiting greater altitudes becomes very important. On the other hand, it could also be possible under altered circumstances, body actively adapted to the required bacterial flora as it is very common during pregnancy[55] . It is quite conceivable that living at HA under multiple harsh conditions like cold, hypoxia, radiations etc., requires

well balanced and strong gut microbiota to support the nutrients availability for metabolism. Nature plays its role by altering the microbial diversity which helps in building an adaptive response to the extreme conditions.

Comparison of genera at different altitudes, revealed the changes within Bacteriodetes, majorly due to significant alteration in the genera Bacteroides & Prevotella. Human gut has been classified into three enterotypes and each one is predominantly occupied by Prevotella, Bacteroides and Ruminococcus because they explain the most human taxonomic variations[56,57,58] . In the present study, the analysis of microbiota at different altitudes revealed Bacteroides enterotypes significantly higher at sea level H1 and moderate at altitude H3 whereas Prevotella enterotypes prevailed at H2, H3 and H4. Our findings are in line with the previous reports of negatively correlated relative abundance of Bacteroides and Prevotella[59] , suggesting their antagonistic nature, that needs to be ascertained.

It is believed that dietary factor has a dominant role in different gut microbiota[23,60] . In the present study the general diet of these sojourners at sea level was low fiber, high protein and high fat, whereas at HA from 3500m onwards the diet was composed of high fiber and low protein. The present study indicates that the anti-correlation may not exclusively be due to diet, because the diet of sojourners was same at H2, H3, and H4, yet the abundance of Bacteroides at H3 was close to that of at H1 than at H2 and H4. Interestingly, at H4 the abundance of Bacteroides was significantly higher than at H2. However the prevalence and abundance of Prevotella was higher at H2, H3 and H4. Our findings on Indian population are not in agreement with the analysis of Li et al[29]reporting that most Han population at 3600 m belonged to the Bacteroides enterotype, rather in agreement with most Tibetans staying at HA at 4800m who are of the Prevotella enterotype. Das et al have also reported Prevotella abundance at 3500m, with highest percentage of individuals consuming non-vegetarian diet[30] . In the present study out of 19 subjects only four were vegetarians, therefore higher abundance of Prevotella in 15 non-vegetarian individuals consuming similar diet at different altitudes was quite interesting. These findings are in contrast to the studies indicating association of Prevotella with high levels of carbohydrates and fruits & vegetables intake[61,62] , whereas in line with other studies confirming with highest percentage of individuals consuming non- fiber diet[63] . Interesting to note that western Indian population at 53m above sea level with highest percentage of individuals consuming carbohydrate and fibrerich components[31] , and north Indian population at 197m with majority of non-vegetarians[30] , have predominance of Prevotella. It is also indicated that omnivorous oligotypes of Prevotella are unique and significantly high, specifically associated with animal origin diet in Indian population[30,23] . Our analysis at four different altitudes revealed association of Prevotella with non-vegetarian dietary pattern which was further confirmed on sub-typing and revealing higher abundance of P11 and P12 oligotypes. The association of these oligotypes has been established with animal origin diet and tri-methyl amine oxide concentration, a carnitine and choline catabolism derived metabolite related to development of CVD[23] . These evidences of different responses of Bacteroides and Prevotella in Indian population clearly indicate the involvement of factors other than diet pointing towards the dominance of Prevotella at sea level[20,30,31] , which increased with the increment in the altitude. Therefore, alteration in the genera Bacteroides and Prevotella cannot be attributed to the dietary or genetic factors only but could be the altitude or hypoxia or other environmental factors which needs to be further looked into.

Studies have reported short term exposure to HA causes coagulation activation[64,65] . Some species of Fusobacterium are known to induce platelet aggregation and portal vein thrombosis[66,67,68] . The source of bacteraemia in these patients is not always evident. In the present study sudden appearance of Fusobacterium at H3 raises alarm suggesting thereby in cases of Fusobacterium bacteraemia of unclear aetiology, portal vein thrombosis should be ruled out at HA considering conditions like low temperature & oxygen pressure, humidity and high radiations etc[67] . Fusobacterium abundance was positively correlated with LDL cholesterol and total cholesterol[69] .

Further, we applied MIDAS for in depth analysis of stool metagenome at different heights and used rare SNPs to track upto strain level abundance to reveal extensive structure and dynamics to elucidate the role of altered gut microbioata in human health and diseases at different altitudes. The approach revealed Prevotella

copri 61740 as the most abundant at HA. With that the door is opened up for further research to understand the role of Prevotella copri 61740 in the pathophysiology of HA diseases.

It is surprising and interesting to note that human gut microbiomes are functionally stable with very little variance in abundance or gene families as "variables". At the same time the relative abundance of host biological pathways is less variable than the relative abundances of microbial phyla in the same metagenome[70,71] . Being the signature of health and disease, strong selection of microbes encodes functions necessary for adaptation to the gut environment indicating the correlation between the two. Thus, implying that human gut microbiota is phylogenetically same, doing similar work but when exposed to extreme conditions, plays differently, demonstrating a natural phenomenon for healing and tailoring therapies. Therefore, in the present study along with knowing composition of the microbial community, DNA sequencing for functional screening using metagenomic approach, was matched with known functional gene sequences. Identifying the genes involved in specific metabolic pathways, could predict functional capabilities, but without messenger RNA, protein and metabolite profiling, these were still predictions. We found significant microbiome alterations that are consistent with well-known side- effects of altitude conditions. The abundances of the predicted pathways emerging as significantly differentiating between four heights, appear to have a pattern of distinct functional divergence that grossly corresponds to environmental conditions. It is reported that the anaerobic environment of gut increases the tolerance capacity to oxidative stress and thus leads to microbial dysbiosis[43] . The significant presence of other bacteria including Roseburia inulinivorans, Faecalibacterium prausnitzii, Oscillibacter and Eubacterium eligens, Eubacterium rectale, Megamonas hypermegale and Prevotella copri which had a very distinctive and increasing appearance at all heights and contrary to the Bacteroides vulgatus, encode butyrate producing functions.

It is suggested that Prevotella is a beneficial bacteria as it is associated with a plant-rich diet, but at the same time due to its high species and strain genetic diversity[72,73] , some of its strains become clinically important pathobionts as they promote and are linked to chronic inflammatory conditions**,** as arthritis, periodontitis, metabolic & cardiovascular disorders and systemic & mucosal T-cell activation in untreated HIV infection[74,75] . Present study is indicative of Prevotella being one of the possible reasons for the gut inflammation at HA. Because of species heterogeneity, it needs to continue the in depth metagenomic analysis of the microbiota in inflammatory diseases so as to reveal its disease modulating role. Immunomodulatory role of Prevotella are reported with significant increase in serum IgA in Rheumatic arthritis which in turn may have implications for disease risk and outcomes[76]**.** Studies are being performed to delve into the basis of its associations with health and disease like effects on host metabolism, providing glucose tolerance in healthy individuals[77] . Reports indicate association between degradation of antibiotics (Xenobiotic degradation) and the glycan metabolism with the abundance of Bacteriodetes[31,78] mainly with abundance of Prevotella, but in the present study glycan metabolism was higher at H3 and lower at H2 and H4 which is converse with the hypothesis. This association between altered gut microbiota at extreme altitude and the functions demonstrate the effectiveness of functional information which could be a beneficial tool for studying disease mechanism rather than a taxonomical marker. It is therefore relevant and important considering the effects of Prevotella on host health under extreme climatic conditions at HA.

Conclusion

Our analyses of database coverage of sojourners belonging to the same ethnicity with similar dietary habits, visiting different altitudes, points to specifically to environmental conditions for changes in the microbial profile, in addition to the sojourners intense physical activities like climbing, excessive temperature fluctuations, isolation, psychological stress, altitude changes, UV exposure, irregular feeding and reduced fluids consumption. The appearance of higher abundance of Prevotella could be one of the causative agents for high incidences of inflammatory diseases at HA, which needs to be revealed in depth. It is imperative to suggest that predicted pathways and correlation with taxonomic abundances need further investigation in Indian population visiting HA. Understanding how the microbiota varies across the population, and correlating the variability with specific microbial functions, is emerging as a component of personalized medicine. Future studies are required to understand the extent of these altered gene level patterns, particularly at strain level

analyses, and targeted treatment of energy metabolism disorders involving altered gut microbiota at HA is need of the hour and of the highest priority for the efforts. Various treatment modalities preferably involved are probiotic intervention, change in the lifestyle, dietary habits and some medications.

Declarations

List of Abbreviation

HA - High Altitude

IBD - Inflammatory Bowel Disease

GIB - Gastro Intestinal Bleeding

DNA - Deoxyribonucleic acids

rRNA - Ribosomal RNA

WGS - Whole Genome Sequencing

MIDAS - Metagenomic Intra-Species Diversity Analysis System

METABAT2 - Metagenome binning with abundance and tetranucleotide frequencies

BMI - Body mass index

PCR - Polymerase Chain Reaction

RNA - Ribonucleic acids

MOCAT2 - Metagenomics Assembly and Gene Prediction Toolkit

RefMG.v1 - Reference marker gene

mOTU.v1 - Metagenomic operational taxonomic units

CSS - Cumulative-sum scaling

SOAP - Short Oligonucleotide Analysis Package

FMAP - Functional Mapping and Analysis Pipeline

OTUs - Operational Taxonomic Units

TSS - Total sum scaling

PcoA - Principal Coordinate Analyses

PERMANOVA - Permutational Multivariate Anova

PCA - Principal Component Analysis

CCA - Canonical Correspondence Analysis

PERMDISP - Permutational Analysis of Multivariate Dispersions

LEfSe - Linear discriminant analysis effect size

FDR - False discovery rate

LDA score - Linear discriminant analysis

CAG - Co-abundance of genes

SNPs - Single-nucleotide polymorphism

FASTA - Fast Alignment

DA - differentially abundant

GDP - Guanosine diphosphate

GMP - Guanosine monophosphate

CVD - Cardiovascular disease

LDL - low-density lipoproteins

HIV - Human immunodeficiency virus

IgA - Immunoglobulin A

UV - Ultra-Violet

Ethics Statement and volunteer information

All expedition participants understood the nature of the study and gave the written consent. The ethics committee of Defence Institute of Physiology and Allied Sciences, Defence Research and Development Organisation, Delhi, India approved all the relevant parameters of the study. The study protocols were in accordance with the Helsinki's approved guidelines.

Consent for Publication: Not applicable

Availability of Data and Material: Sequencing data reported in this paper is available at NCBI under the project no. PRJNA492714.

Funding

This work is supported by Defence Research and Development Organization.

Authors Contribution

LG, SBS and BK devised the project, designed the study protocol and supervised all phases of the project. BB, APY and MPKR performed sample collection and DNA extraction. LG and BB carried out subjects phenotyping, data interpretation & analyses, and wrote the manuscript. MPKR and TN arranged for subjects availability and provided logistics support at High Altitude. AKS participated in sample collection at base line, DNA extraction and library preparations. All authors contributed to data interpretation, discussions and editing of the paper.

Acknowledgment: Authors thank DRDO for funding. Also, Nxgenbio is greatly acknowledged for data assembly and analysis. All the sojourns are accredited for participating in the study.

Competing Interests: No authors report any conflict of interest.

References

1. León-Velarde, F., Maggiorini, M., Reeves, J.T., Aldashev, A., Asmus, I., Bernardi, L. et al. (2005). Consensus statement on chronic and subacute high-altitude diseases. High Alt. Med. Biol. *6* , 147-157.
2. Luks, A.M. (2015). Physiology in Medicine: A physiologic approach to prevention and treatment of acute high-altitude illnesses. J. Appl. Physiol. *118* , 509–519.
3. Hooper, L.V., Littman, D.R., and Macpherson, A.J. (2012). Interactions between the microbiota and the immune system. Science *336* , 1268-1273.
4. Cho, I., and Blaser, M. J. (2012). The human microbiome: at the interface of health and disease. Nature Rev. Genet.*13* **,** 260–270.
5. Huttenhower, C., Kostic, A.D., and Xavier, R.J. (2014). Inflammatory bowel disease as a model for translating the microbiome. Immunity*40* , 843-854.
6. Shreiner, A. B., Kao, J. Y., and Young, V. B. (2015). The gut microbiome in health and in disease. Curr. Opin. Gastroenterol. *31* **,** 69–75.

7. Schroeder, B.O., and Bäckhed, F. (2016). Signals from the gut microbiota to distant organs in physiology and disease. Nature Med.*22* , 1079-1089.
8. Shreiner, A. B., Kao, J. Y., and Young, V. B. (2015). The gut microbiome in health and in disease. Curr. Opin. Gastroenterol. *31* **,** 69–75.
9. Cho, I., and Blaser, M. J. (2012). The human microbiome: at the interface of health and disease. Nature Rev. Genet.*13* **,** 260–270.
10. Hooper, L.V., Littman, D.R., and Macpherson, A.J. (2012). Interactions between the microbiota and the immune system. Science *336* , 1268-1273.
11. Huttenhower, C., Kostic, A.D., and Xavier, R.J. (2014). Inflammatory bowel disease as a model for translating the microbiome. Immunity*40* , 843-854.
12. Schroeder, B.O., and Bäckhed, F. (2016). Signals from the gut microbiota to distant organs in physiology and disease. Nature Med.*22* , 1079-1089.
13. Meer, R.R., Songer, J.G., and Park, D.L. (1997). Human disease associated with Clostridium perfringens enterotoxin. Rev. Environ. Contam. Toxicol. *150* , 75-94.
14. Veldhuyzen van Zanten, S.J., Pollak, P.T., Best, L.M., Bezanson, G.S. and Marrie, T. (1994). Increasing prevalence of Helicobacter pylori infection with age: continuous risk of infection in adults rather than cohort effect. Journal of Infectious Diseases *169* ,434-437.
15. Tong, S.Y., Davis, J.S., Eichenberger, E., Holland, T.L., and Fowler, V.G. Jr. (2015). Staphylococcus aureus infections: epidemiology, pathophysiology, clinical manifestations, and management. Clin. Microbiol. Rev. *28* , 603-661.
16. Adak, A., Maity, C., Ghosh, K., and Mondal, K.C. (2014). Alteration of predominant gastrointestinal flora and oxidative damage of large intestine under simulated hypobaric hypoxia. Z Gastroenterol.*52* , 180-186.
17. Adak, A., Ghosh, K., and Mondal, K.C. (2014). Modulation of small intestinal homeostasis along with its microflora during acclimatization at simulated hypobaric hypoxia. Ind. J. Exp. Biol.*52* , 1098-1105.
18. Berrios, J. (1982). Consideraciones sobre la pathologia digestive en los habitentes de las grandes alturasdel Peru. (In Spanish). Rev Gastroenterol (Peru), pp. 21-28.
19. Wu, T.Y. (2001). Take note of altitude gastrointestinal bleeding. Newsletter Int. Soc. Mountain Med. *10* , 9-11.
20. Saito, A. (1989). The medical reports of the China-Japan-Nepal Friendship Expedition to Mt. Qomolungma/Sagamartha (Everest). Jap. J. Mount. Med. 9, 83-87.
21. Liu, M.F. (1995). Upper alimentary bleeding at high altitude. In: Lu, Y.D., Li, K.X., Yin, Z.Y., editors. High Altitude Medicine and Physiology. (In Chinese). Tianjing: Tianjing Science & Technology Press, 586.
22. Ley, R.E., Peterson, D.A., and Gordon, J.I. (2006). Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell *24* , 837-848.
23. Sekirov, I., Russell, S.L., Antunes, L.C., and Finlay, B.B. (2010). Gut microbiota in health and disease. Physiol. Rev. *90* , 859-904.
24. Bhute, S., Pande, P., Shetty, S.A., Shelar, R., Mane, S., Kumbhare, S.V. et al. (2016). Molecular characterization and meta-analysis of gut microbial communities illustrate enrichment of Prevotella and Megasphaera in Indian subjects. Front. Microbiol. *9* , 657-660.
25. Dehingia, M., Devi, K.T., Talukdar, N.C., Talukdar, R., Reddy, N., Mande, S.S. et al. (2015). Gut bacterial diversity of the tribes of India and comparison with the worldwide data. Sci. Rep. *22* , 18555-18563.
26. Ghosh, T.S., Gupta, S.S., Bhattacharya, T., Yadav, D., Barik, A., Chowdhury, A. et al. (2014). Gut microbiomes of Indian children of varying nutritional status. PLoS One *9* , e95547.
27. Dethlefsen, L., and Relman, D.A. (2011). Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. Proc. Natl. Acad. Sci. *108* , 4554-4561.
28. De Filippo, C., Cavalieri, D., Di Paola, M., Ramazzotti, M., Poullet, J.B., Massart, S. et al. (2010). Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and

rural Africa. Proc. Natl. Acad. Sci. *107* , 14691-14696.
29. Khachatryan, Z.A., Ktsoyan, Z.A., Manukyan, G.P., Kelly, D., Ghazaryan, K.A., and Aminov, R.I. (2008). Predominant role of host genetics in controlling the composition of gut microbiota. PLoS One*3* , e3064.
30. Goodrich, J.K., Waters, J.L., Poole, A.C., Sutter, J.L., Koren, O., Blekhman, R. et al. (2014). Human genetics shape the gut microbiome. Cell *159* , 789-799.
31. Kleessen, B., Schroedl, W., Stueck, M., Richter, A., Rieck, O., and Krueger, M. (2005). Microbial and immunological responses relative to high-altitude exposure in mountaineers. Med. Sci. Sports Exerc.*37* , 1313-1318.
32. Adak, A., Maity, C., Ghosh, K., Pati, B.R., and Mondal, K.C. (2013). Dynamics of predominant microbiota in the human gastrointestinal tract and change in luminal enzymes and immunoglobulin profile during high-altitude adaptation. Folia. Microbiol. (Praha) *58* , 523-528.
33. Li, L., and Zhao, X. (2015). Comparative analyses of fecal microbiota in Tibetan and Chinese Han living at low or high altitude by barcoded 454 pyrosequencing. Sci Rep. *5* , 1-10.
34. Li, K., Dan, Z., Gesang, L., Wang, H., Zhou, Y., Du, Y. et al. (2016). Comparative Analysis of Gut Microbiota of Native Tibetan and Han Populations Living at Different Altitudes. PLoS One *11* , e0155863.
35. Das, B., Ghosh, T.S., Kedia, S., Rampal, R., Saxena, S., Bag, S. et al. (2018). Analysis of the Gut Microbiome of Rural and Urban Healthy Indians Living in Sea Level and High Altitude Areas. Sci. Rep.*8* , 1-15.
36. Tandon, D., Haque, M.M., Saravanan, R., Shaikh, S., Sriram, P., Dubey, A.K. et al. (2018). A snapshot of gut microbiota of an adult urban population from Western region of India. PLoS One *13* , e0195643.
37. Lepage, P., Leclerc, M.C., Joossens, M., Mondot, S., Blottière, H.M., Raes, J. et al. (2013). A metagenomic insight into our gut's microbiome. Gut *62* , 146-158.
38. Lieberman, T.D., Flett, K.B., Yelin, I., Martin, T.R., McAdam, A.J., Priebe, G.P. et al. (2013). Genetic variation of a bacterial pathogen within individuals with cystic fibrosis provides a record of selective pressures. Nature Genet. *46* , 82-87.
39. Snitkin, E.S., Zelazny, A.M., Montero, C.I., Stock, F., Mijares, L., NISC Comparative Sequence Program. et al. (2011) Genome-wide recombination drives diversification of epidemic strains of Acinetobacter baumannii. Proc. Natl. Acad. Sci. *108* , 13758-13763.
40. Rasko, D.A., Webster, D.R., Sahl, J.W., Bashir, A., Boisen, N., Scheutz, F. et al. (2011). Origins of the E. coli strain causing an outbreak of hemolytic-uremic syndrome in Germany. N. Engl. J. Med.*365* , 709-717.
41. Lauber, C.L., Zhou, N., Gordon, J.I., Knight, R., and Fierer, N. (2010). Effect of storage conditions on the assessment of bacterial community structure in soil and human-associated samples. FEMS Microbiol. Lett. *307* , 80-86.
42. Carroll, I.M., Ringel-Kulka, T., Siddle, J.P., Klaenhammer, T.R., and Ringel, Y. (2012). Characterization of the fecal microbiota using high-throughput sequencing reveals a stable microbial community during storage. PLoS One *7* , e46953.
43. Kultima, J.R., Coelho, L.P., Forslund, K., Huerta-Cepas, J., Li, S.S., Driessen, M. et al. (2016). MOCAT2: a metagenomic assembly, annotation and profiling framework. Bioinformatics *32* , 2520-2523.
44. Mende, D.R., Sunagawa, S., Zeller, G., and Bork, P. (2013). Accurate and universal delineation of prokaryotic species. Nat. Met. *10* , 881-884. doi: 10.1038/nmeth.2575.
45. Paulson, J.N., Stine, O.C., Bravo, H.C., and Pop, M. (2013). Differential abundance analysis for microbial marker-gene surveys. Nat. Met. *10* , 1200-1202.
46. Zakrzewski, M., Proietti, C., Ellis, J.J., Hasan, S., Brion, M.J., Berger, B. et al. (2017). Calypso: a user-friendly web-server for mining and visualizing microbiome-environment interactions. Bioinformatics *1* , 782-783.
47. Kang, D.D., Froula, J., Egan, R., and Wang, Z. (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. Peer J. *3* , e1165.
48. Nayfach, S., Rodriguez-Mueller, B., Garud, N., and Pollard, K.S. (2016). An integrated metagenomics

pipeline for strain profiling reveals novel patterns of bacterial transmission and biogeography. Genome Res. *26* , 1612-1625.

49. Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J. et al. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. Giga Science 1, 1-6.
50. Medini, D., Donati, C., Tettelin, H., Masignani, V., and Rappuoli, R. (2005). The microbial pan-genome. Curr. Opin. Genet. Dev. *15* , 589-594.
51. Jensen, L.J., Julien, P., Kuhn, M., von Mering, C., Muller, J., Doerks, T. et al. (2008). eggNOG: automated construction and annotation of orthologous groups of genes. Nucleic Acids Res.*36* , D250-D254.
52. Kim, J., Kim, M.S., Koh, A.Y., Xie, Y., and Zhan, X. (2016). FMAP: Functional Mapping and Analysis Pipeline for metagenomics and meta-transcriptomics studies. BMC Bioinformatics *17* , 1-8.
53. De Filippis, F., Pellegrini, N., Laghi, L., Gobbetti, M., and Ercolini, D. (2016). Unusual sub-genus associations of faecal Prevotella and Bacteroides with specific dietary patterns. Microbiome*4* , 1-6.
54. Khatri, P., Sirota, M., and Butte, A.J. (2012). Ten years of pathway analysis: current approaches and outstanding challenges. PLoS Comput. Biol. *8* , e1002375.
55. Vogtmann, E., Hua, X., Zeller, G., Sunagawa, S., Voigt, A.Y., Hercog, R. et al. (2016). Colorectal Cancer and the Human Gut Microbiome: Reproducibility with Whole-Genome Shotgun Sequencing. PLoS One*11* , e0155362.
56. Donaldson, G. P., Lee, S. M., and Mazmanian, S. K. (2016). Gut biogeography of the bacterial microbiota. Nature Rev. Microbiol. *14* , 1-13.
57. Litvak. Y., Sharon, S., Hyams, M., Zhang, L., Kobi, S., Katsowich, N. et al. (2017). Epithelial cells detect functional type III secretion system of entero pathogenic Escherichia coli through a novel NF-ϰB signaling pathway. PLoS Pathog. *13* , e1006472.
58. WHO E.C. (2004). Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet*363* , 157-63.
59. Nemati, R., Dietz, C., Anstadt, E.J., Cervantes, J., Liu, Y., Dewhirst, F.E. et al. (2007). Deposition and hydrolysis of serine dipeptide lipids of Bacteroidetes bacteria in human arteries: relationship to atherosclerosis. J. Lipid Res. *58* , 1999-2007.
60. Konstantinov, S.R., van der Woude, C.J., and Peppelenbosch, M.P. (2013). Do pregnancy-related changes in the microbiome stimulate innate immunity? Trends. Mol. Med. *19* , 454-459.
61. Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D.R. et al. (2011). Enterotypes of the human gut microbiome. Nature *473* , 174-180.
62. Vandeputte, D., Falony, G., Vieira-Silva, S., Tito, R.Y., Joossens, M., and Raes, J. (2016). Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. Gut *65* , 57-62.
63. Falony, G., Joossens, M., Vieira-Silva, S., Wang, J., Darzi, Y., Faust, K. et al. (2016). Population-level analysis of gut microbiome variation. Science *352* , 560-564.
64. Moeller, A.H., Li, Y., Ngole, E.M., Ahuka-Mundeke, S., Lonsdorf, E.V., Pusey, A.E. et al. (2014). Rapid changes in the gut microbiome during human evolution. Proc. Natl. Acad. Sci. *111* , 16431-16435.
65. Conlon, M. and Bird, A., 2014. The impact of diet and lifestyle on gut microbiota and human health. Nutrients *7* , 17-44.
66. De Filippis, F., Pellegrini, N., Vannini, L., Jeffery, I.B., La Storia, A., Laghi, L. et al. (2016). High-level adherence to a Mediterranean diet beneficially impacts the gut microbiota and associated metabolome. Gut *65,* 1812-1821.
67. Wu, G.D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y.Y., Keilbaugh, S.A. et al. (2011). Linking long-term dietary patterns with gut microbial enterotypes. Science *334* ,105-108.
68. McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., DeSantis, T.Z., Probst, A. et al. (2011). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J. *6* , 610-618.
69. Mannucci, P.M., Gringeri, A., Peyvandi, F., Di Paolantonio, T., and Mariani, G. (2002). Short-term

exposure to high altitude causes coagulation activation and inhibits fibrinolysis. Thromb. Haemost. *87* , 342-343.

70. Tyagi, T., Ahmad, S., Gupta, N., Sahu, A., Ahmad, Y., Nair, V. et al. (2014). Altered expression of platelet proteins and calpain activity mediate hypoxia-induced prothrombotic phenotype. Blood *123* , 1250-1260.
71. Bultink, I.E., Dorigo-Zetsma, J.W., Koopman, M.G., and Kuijper, E.J. (1999). Fusobacterium nucleatum septicemia and portal vein thrombosis. Clin. Infect. Dis. *28* , 1325-1326.
72. Hamidi, K., Pauwels, A., Bingen, M., Simo, A.C., Medini, A., Jarjous, N. et al. (2008). Recent portal and mesenteric venous thrombosis associated with Fusobacterium bacteremia. Gastroenterol. Clin. Biol.*32* , 734-739.
73. Redford, M.R., Ellis, R., and Rees, C.J. (2005). Fusobacterium necrophorum infection associated with portal vein thrombosis. J. Med. Microbiol. *54* , 993-995.
74. Koren, O., Spor, A., Felin, J., Fak, F., Stombaugh, J., Tremaroli, V. et al. (2011). Human, oral, gut, and plaque microbiota in patients with atherosclerosis. Proc. Natl. Acad. Sci. *108* , 4592-4598.
75. Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley, R.E. et al. (2009). A core gut microbiome in obese and lean twins. Nature *457* , 480–484.
76. Human Microbiome Project Consortium (2012). Structure, function and diversity of the healthy human microbiome. Nature *486* , 207–214.
77. Zhu, A., Sunagawa, S., Mende, D.R., and Bork, P. (2015). Inter-individual differences in the gene content of human gut bacterial species. Genome Biol. *16* , 1-13.
78. Ley, R.E., Peterson, D.A., and Gordon, J.I. (2006). Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell *24* , 837-848.
79. Scher, J.U., Sczesnak, A., Longman, R.S., Segata, N., Ubeda, C., Bielski, C. et al. (2013). Expansion of intestinal Prevotella copri correlates with enhanced susceptibility to arthritis. eLife*2* , 1-20.
80. Dillon, S.M., Lee, E.J., Kotter, C.V., Austin, G.L., Gianella, S., Siewe, B. et al. (2016). Gut dendritic cell activation links an altered colonic microbiome to mucosal and systemic T-cell activation in untreated HIV-1 infection. Mucosal Immunol. *9* , 24-37.
81. Larsen, J.M. (2017). The immune response to Prevotella bacteria in chronic inflammatory disease. Immunology *151* , 363-374.
82. Kovatcheva-Datchary, P., Nilsson, A., Akrami, R., Lee, Y.S., De Vadder, F., Arora, T. et al. (2015). Dietary Fiber-Induced Improvement in Glucose Metabolism Is Associated with Increased Abundance of Prevotella. Cell Metab. *22* , 971-982.
83. Maurice, C.F., Haiser, H.J., and Turnbaugh, P.J. (2013). Xenobiotics shape the physiology and gene expression of the active human gut microbiome. Cell *152* , 39-50.
84. Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K. et al. (2010). QIIME allows analysis of high-throughput community sequencing data. Nat. Met. *7* , 335-336.

Figure legends

Figure 1: ***(A)*** *Differences in bacterial alpha diversity (Shannon diversity, Richness, Evenness Chao and Fishers Alpha) between H1 & H2 heights.* ***(B)*** *Differences in bacterial alpha diversity (Shannon diversity, Richness, Evenness, Chao and Fishers Alpha at all heights (H1, H2, H3 & H4).* ***(C)*** *Rarefraction curvewith error bars showing confidence intervals, for all four groups of samples based on OTUs detected using a similarity threshold of 97% (0.97) and phylum and genus levels.*

Figure 2: *Supervised PCA and Canonical Correspondence Analysis (CCA) displaying the compositional distribution of the fecal microbiota at altitudes H1 and H2, comparing 16s rRNA and WGS results:* ***(Ai)*** *16srRNA PCA and* ***(Aii)*** *CCA analysis,* ***(Bi)****WGS PCA and* ***(Bii)*** *CCA analysis. PERMDISP 2 Analysis indicating differences in global microbial community composition comparison at two heights H1 and H2:* ***(Aiii)*** *16srRNA and* ***(Biii)*** *WGS. PERMDISP 2 Analysis indicating differences in global microbial community composition comparison at all heights H1, H2, H3, and H4* ***(C)****PERMDISP 2,* ***(D)*** *CCA analysis*
.

Figure 3: ***(A)*** *Differentially abundant & significant taxa in WGS data at phylum level (evaluated using T test at a p-value cut off of 0.05) at four heights H1, H2, H3 and H4.****(B)*** *Odds ratio at 95 % CI &at a p-value cut off of 0.05 at Heights H2 vs H1.* ***(C)*** *Differentially abundant & significant taxa in 16SrRNA data at Genus level using Wilcoxon rank test****(D)*** *Q-Q plots of p values for the association between the relative abundance and presence/absence of phyla.*

Figure 4: *(A) Differentially abundant & significant taxa in WGS data at Genus level (evaluated using T test at a p-value cut off of 0.05) at four heights H1, H2, H3 and H4. (B) Odds ratio at 95 % CI &at a p-value cut off of 0.05 at Heights H2 vs H1. (C) Differentially abundant & significant taxa in 16SrRNA data at Genus level using Wilcoxon rank test (D) Q-Q plots of p values for the association between the relative abundance and presence/absence of Genus.*

Figure 5*: FastTree uses, generalised time-reversible (GTR) models of nucleotide evolution in addition to the CAT Model (evolutionary rate heterogeneity) approximation, and finally for estimating the reliability of each split in the tree using the Shimodaira-Hasegawa test, at 1000 resamples.*

Figure 6: *Individual Significant Bacterial Profiles at all Time points.*

Figure 7*:* ***(A)*** *Functional characteristics of gut microbiomes of individuals at different altitudes (H2, H3 & H4) as compare with H1.* ***(B)*** *Alteration in metabolic pathways at H1 and H2 predicted by 16sRNA.*