Metagenome of the gut of a malnourished child

Sourav Sen Gupta1, Monzoorul Haque Mohammed2, Tarini Shankar Ghosh2, Suman Kanungo1, Gopinath Balakrish Nair1 and Sharmila S Mande2*

Abstract

Background: Malnutrition, a major health problem, affects a significant proportion of preschool children in developing countries. The devastating consequences of malnutrition include diarrhoea, malabsorption, increased intestinal permeability, suboptimal immune response, etc. Nutritional interventions and dietary solutions have not been effective for treatment of malnutrition till date. Metagenomic procedures allow one to access the complex cross-talk between the gut and its microbial flora and understand how a different community composition affects various states of human health. In this study, a metagenomic approach was employed for analysing the differences between gut microbial communities obtained from a malnourished and an apparently healthy child.

Results: Our results indicate that the malnourished child gut has an abundance of enteric pathogens which are known to cause intestinal inflammation resulting in malabsorption of nutrients. We also identified a few functional sub-systems from these pathogens, which probably impact the overall metabolic capabilities of the malnourished child gut.

Conclusion: The present study comprehensively characterizes the microbial community resident in the gut of a malnourished child. This study has attempted to extend the understanding of the basis of malnutrition beyond nutrition deprivation.

Background

Malnutrition is a major global problem. While one form of malnutrition (referred to as under-nutrition) encompasses stunting, wasting, and deficiencies of essential vitamins and minerals, the other form manifests as obesity due to over-consumption of specific nutrients. The prevalence of underweight, stunting, and wasting in children, the most reliable measures of malnutrition, is concentrated in few countries in South Asia and Eastern Africa [1,2] where 33% and 28%, respectively, of the children younger than 5 years are underweight. Despite its rapid economic growth in the last decade, the estimated prevalence of child stunting in India is 51%. This translates to approximately 61 million stunted children, constituting 34% of the global total. Efforts to reduce the proportion of underweight children by half by the year 2015 is set as the first millennium development goal (MDG-1) by the United Nations.

The consequences of malnutrition are devastating. These include diarrhoea, malabsorption, small bowel overgrowth, increased intestinal permeability, enteropathy, gram-negative (enteric) bacteraemia, and suboptimal immune response [3,4]. Nutritional interventions and dietary solutions have not been the most effective till date for treatment of malnutrition. However, recent literature sheds some light on our simplistic understanding of the basis of malnutrition. Composition of host gut microbiome has been thought to play an extremely important role in absorption of nutrients from food and response to caloric deficit. Equal access to calories is now thought to be not limiting in the establishment of malnutrition. Even enteric infections result in malabsorption of nutrients due to intestinal inflammation and contribute to malnutrition in large proportions of children in developing countries. The genomes of large number of microbes present in the human gut endow us with physiological capabilities that we have not had to evolve on our own and contribute immensely in manifestation of who we are genetically and metabolically, and a reflection of our state of well being.

Recent development of metagenomic procedures has now enabled us to access the complex cross-talk between the gut and its microbial flora and understand how a different

* Correspondence: sharmila.mande@tcs.com
2Bio-Sciences R&D Division, TCS Innovation Labs Hyderabad, Tata Consultancy Services Limited, Hyderabad, India
Full list of author information is available at the end of the article

c© 2011 Gupta et al; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
community composition affects various states of human health.

In the present study, we examined the gut microbiota using faecal samples from a malnourished child and another apparently healthy child as a control from a typical urban slum setting in Kolkata, India and attempted to identify compositional and functional differences as derived from the two metagenomes. We anticipated that by analysing the gut microbial communities from the two subjects, we could have an understanding of the underlying influence of bacterial inhabitants of the gut in malnourished and healthy conditions. The overall objective was to understand the relationship between the nutritional status and the microbial community in the gut.

Results and Discussion

Pyrosequencing of metagenomes obtained from malnourished and healthy child faecal DNA samples yielded 14,96,170 and 12,71,252 high-quality sequence reads, respectively. Eu-Detect analysis (http://metagenomics.atic.tcs.com/Eu-Detect/) followed by subsequent BLASTn searches against the human genome identified 10.3% of the sequences in the malnourished and 0.3% sequences in the healthy child data sets to have originated from human DNA. This indicates a probable exfoliation of human tissues, manifested as contamination in the sequenced faecal sample obtained from the malnourished child.

A total of 41.1% and 37% of reads in malnourished and healthy child data sets, respectively, could be classified under various taxonomic groups using SPHINX algorithm [5]. Results indicated a striking abundance of four bacterial lineages in the gut of the malnourished child as compared to the healthy child (Figure 1, Table 1).

Families Campylobacteraceae and Helicobacteraceae were 35 and 12 folds higher in the malnourished child sample (Figure 1) suggesting infection of the intestinal epithelium by gastrointestinal pathogens belonging to these families. It is well known that infection adversely affects nutritional status and also that malnutrition can predispose to infection [6]. Though it is difficult to predict the exact sequence of events in the vicious cycle of infection and malnutrition, the present study underscores the presence of pathogens normally absent in the intestines of healthy subjects. Species belonging to Campylobacter and Helicobacter are well-established human pathogens. For example, Campylobacter jejuni and Campylobacter coli are known to be the causative agents of Campylobacteriosis, a collective description for diarrhoeal disease caused by members of the Campylobacter genus. Similarly, the microaerophilic bacterium Helicobacter, known to inhabit various areas of the stomach (particularly the antrum), causes a chronic low-level inflammation of the stomach lining and is linked to the development of duodenal and gastric ulcers and stomach cancer. An association between reduction in gastric acid and Helicobacter pylori infections has been linked to increase in Vibrio cholerae infections [7,8]. Individuals with gastric hypochlorhydria or achlorhydria are at greater risk of developing cholera after infection with low inoculums [9]. The high density of members of the family Helicobacteraceae in the malnourished child gut metagenome would predispose this child to repeated infections because gastric acid is known to act as a natural non-specific barrier to enteric pathogens. Therefore, in the malnourished children, a lower magnitude of bacterial inoculum would give rise to severe infection as compared to their healthy counterparts. In the healthy child gut metagenome, such a predilection of Helicobacteraceae sequences was not observed suggesting that the healthy flora could competently inhibit and exclude extraneous pathogens. A malnourished child, in contrast, would suffer from repeated bouts of infections (and disease) and this coupled with poor nutrient absorption would drive the child into throes of declining health and ultimately to death.

Bacteroidaceae family was also observed to have an appreciably higher representation in the malnourished sample (Figure 1). In contrast to 3% of sequences in the healthy child sample, approximately 13% of sequences belonged to this family in the malnourished child sample. An increase in abundance of this bacterial phylotype has been previously associated with a decrease in body weight of obese individuals [10]. Similarly, the family Porphyromonadaceae was also seen to have approximately a two-fold enrichment in the malnourished child sample. Interestingly, members of the Porphyromonadaceae family were shown to be exclusively present in the faecal microbiota of patients suffering from Crohn’s disease, an inflammatory bowel disease [11].

To understand the functional differences between these two gut metagenomes, we used HabiSign algorithm (http://metagenomics.atic.tcs.com/HabiSign/) to identify sequences unique to either healthy or malnourished samples. The taxonomic affiliations of these sequences were inferred before functionally analysing them. 79% of the sequences specific to the malnourished child sample belonged to the order Campylobacterales (Figure 2, Table 2). Another subset of malnourished sample specific sequences belonged to the order Clostridiales. Surprisingly, this order was observed to be equally represented in both samples by the SPHINX-based analysis (Table 1). This observation indicates the presence of novel species belonging to the order Clostridiales in the malnourished child metagenome.

In contrast, sequences specific to the healthy child data set were seen to be enriched for the following orders - Lactobacillales, Enterobacteriales, Pseudomonadales,
Chloroflexales, Xanthomonadales, Planctomycetales, Halobacterales, Burkholderiales, Actinomycetales, Bifidobacteriales, Desulfovibrionales and Rhizobiales (Figure 2, Table 2). The microbial fitness of the healthy child gut in performing normal gut function is enhanced due to the overall abundance of these bacterial phylotypes and absence of potentially harmful enteropathogens. A thriving gut bacterial community in the healthy child also helps in the proper maturation of immunity which in turn aids in defence against gut pathogens. As a result of reduced abundance of many beneficial members of the known bacterial phylotypes in the malnourished gut as compared to the healthy child, there is also a reduced availability of usable microbial fermented products from otherwise indigestible dietary polysaccharides and microbe derived nutrients in the malnourished gut.

Analysis of the results obtained from the Meta Genome Rapid Annotation using Subsystem Technology (MG-RAST) server [12] for the identified malnourished/healthy metagenome specific sequences indicated the presence of functionally characterized protein encoding genes (PEGs) that were specifically abundant in either sample (Table 3). Results indicated that PEGs belonging to four subsystems (motility and chemotaxis, respiration, membrane transport, virulence) were relatively more abundant in sequences identified as specific to the malnourished child sample (Figure 3).

Flagellum, flagellum motility subsystems and motility accessory proteins, all mapping to species belonging to the order Campylobacterales, were abundant in the malnourished child sample under the motility and chemotaxis subsystem (Figure 3). Similarly, approximately 70% of malnourished specific sequences functionally classified under the ‘respiration’ category were also seen to be taxonomically assigned to the order Campylobacterales (Figure 3).

Malnourished specific sequences tagged to the ‘membrane transport’ subsystem also indicated an abundance (~70%) of sequences classified under the order Campylobacterales (Figure 3). While some of these sequences mapped to ABC transporters of branched-chain amino acids, Zinc, Tungstate and Molybdenum, others mapped to specialized membrane transport systems like Na(+)-H (+) antiporter and the lipoprotein sorting systems. Specific Tungstate and Molybdenum transport systems are known to exist in pathogenic Campylobacter species [13,14]. Lipoproteins are known to play a major role in the virulence of several pathogenic organisms. Besides
having key roles in functional pathways like motility, chemotaxis, cell-cell interactions, and signal transduction, lipoproteins are also implicated in the assembly and regulation of bacterial secretion systems [15].

Many malnourished child specific virulence-associated sequences belonged to order Campylobacterales (Figure 3). These sequences mapped to PEGs corresponding to iron metabolism, adhesion, multidrug resistance efflux pumps, hemin transport system, iron transport, Ton and Tol systems. 31% of virulence-associated malnourished specific sequences, mapping to the above PEGs, could only be assigned at the superkingdom level of Bacteria (Figure 3), indicating presence of hitherto unknown organism(s) with pathogenic potential in the malnourished child gut. Interestingly, 7.6% of the virulence-associated sequences mapped to the PEGs associated with various components

Table 1 Comparison of the taxonomic assignments obtained (using SPHINX algorithm) for malnourished and healthy samples at the taxonomic levels of family, order, class and phylum

<table>
<thead>
<tr>
<th>Taxonomic Level</th>
<th>Taxon Name</th>
<th>% of sequences assigned</th>
<th>Malnourished Sample (X)</th>
<th>Healthy Sample (Y)</th>
<th>Relative Ratio (X/Y)</th>
<th>Inference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Campylobacteraceae</td>
<td>16.7</td>
<td>0.5</td>
<td>35.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Helicobacteraceae</td>
<td>12</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Bacteroidaceae</td>
<td>12.8</td>
<td>3</td>
<td>4.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Porphyromonadaceae</td>
<td>5</td>
<td>2.8</td>
<td>1.7</td>
</tr>
<tr>
<td>Family</td>
<td>Clostridaceae</td>
<td>5.9</td>
<td>4</td>
<td>1.5</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bacillaceae</td>
<td>3.7</td>
<td>4.9</td>
<td>0.8</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Staphylococcaceae</td>
<td>1.3</td>
<td>2.1</td>
<td>0.6</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Family</td>
<td>Streptococcaceae</td>
<td>6.1</td>
<td>10.4</td>
<td>0.6</td>
<td>-</td>
<td>High in healthy</td>
</tr>
<tr>
<td></td>
<td>Enterobacteriaceae</td>
<td>12.2</td>
<td>22.5</td>
<td>0.5</td>
<td>-</td>
<td>High in healthy</td>
</tr>
<tr>
<td></td>
<td>Methanosarcinaceae</td>
<td>1</td>
<td>2.2</td>
<td>0.5</td>
<td>-</td>
<td>High in healthy</td>
</tr>
<tr>
<td>Family</td>
<td>Thermotogaceae</td>
<td>1.1</td>
<td>2.6</td>
<td>0.4</td>
<td>-</td>
<td>High in healthy</td>
</tr>
<tr>
<td></td>
<td>Shewanellaceae</td>
<td>1.6</td>
<td>4.1</td>
<td>0.4</td>
<td>-</td>
<td>High in healthy</td>
</tr>
<tr>
<td></td>
<td>Eubacteriaceae</td>
<td>1.9</td>
<td>5.1</td>
<td>0.4</td>
<td>-</td>
<td>High in healthy</td>
</tr>
<tr>
<td></td>
<td>Campylobacterales</td>
<td>26.64</td>
<td>1.32</td>
<td>20.25</td>
<td>-</td>
<td>High in malnourished</td>
</tr>
<tr>
<td></td>
<td>Bacteroidales</td>
<td>17.42</td>
<td>6.33</td>
<td>2.75</td>
<td>-</td>
<td>High in malnourished</td>
</tr>
<tr>
<td></td>
<td>Clostridales</td>
<td>7.09</td>
<td>8.36</td>
<td>0.85</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bacillales</td>
<td>10.09</td>
<td>12.84</td>
<td>0.79</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Order</td>
<td>Lactobacillales</td>
<td>6.13</td>
<td>10.39</td>
<td>0.59</td>
<td>-</td>
<td>High in healthy</td>
</tr>
<tr>
<td></td>
<td>Enterobacteriales</td>
<td>10.71</td>
<td>19.09</td>
<td>0.56</td>
<td>-</td>
<td>High in healthy</td>
</tr>
<tr>
<td></td>
<td>Methanosarciniales</td>
<td>1.43</td>
<td>2.63</td>
<td>0.55</td>
<td>-</td>
<td>High in healthy</td>
</tr>
<tr>
<td></td>
<td>Alteromonadiales</td>
<td>1.98</td>
<td>4.3</td>
<td>0.46</td>
<td>-</td>
<td>High in healthy</td>
</tr>
<tr>
<td></td>
<td>Thermotogales</td>
<td>0.89</td>
<td>2.08</td>
<td>0.43</td>
<td>-</td>
<td>High in healthy</td>
</tr>
<tr>
<td></td>
<td>Actinomycetaceae</td>
<td>0.54</td>
<td>2.2</td>
<td>0.25</td>
<td>-</td>
<td>High in healthy</td>
</tr>
<tr>
<td>Class</td>
<td>Epsilonproteobacteria</td>
<td>21.99</td>
<td>1.06</td>
<td>20.79</td>
<td>-</td>
<td>High in malnourished</td>
</tr>
<tr>
<td></td>
<td>Bacteroidia</td>
<td>14.25</td>
<td>4.51</td>
<td>3.16</td>
<td>-</td>
<td>High in malnourished</td>
</tr>
<tr>
<td></td>
<td>Clostridia</td>
<td>6.46</td>
<td>6.6</td>
<td>0.98</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bacilli</td>
<td>16.77</td>
<td>20.49</td>
<td>0.82</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Class</td>
<td>Alphaproteobacteria</td>
<td>1.75</td>
<td>2.43</td>
<td>0.72</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Methanomicrobia</td>
<td>1.88</td>
<td>2.96</td>
<td>0.64</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Class</td>
<td>Gammaproteobacteria</td>
<td>24.28</td>
<td>41.37</td>
<td>0.59</td>
<td>-</td>
<td>High in healthy</td>
</tr>
<tr>
<td></td>
<td>Thermoprotei</td>
<td>1.21</td>
<td>2.09</td>
<td>0.58</td>
<td>-</td>
<td>High in healthy</td>
</tr>
<tr>
<td></td>
<td>Actinobacteria (class)</td>
<td>0.61</td>
<td>2.49</td>
<td>0.24</td>
<td>-</td>
<td>High in healthy</td>
</tr>
<tr>
<td>Phylum</td>
<td>Bacteroidetes</td>
<td>13.95</td>
<td>4.88</td>
<td>2.86</td>
<td>-</td>
<td>High in malnourished</td>
</tr>
<tr>
<td></td>
<td>Proteobacteria</td>
<td>49.78</td>
<td>50.61</td>
<td>0.98</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Firmicutes</td>
<td>24.28</td>
<td>26.27</td>
<td>0.92</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Euryarchaeota</td>
<td>3.95</td>
<td>7.03</td>
<td>0.56</td>
<td>-</td>
<td>High in healthy</td>
</tr>
</tbody>
</table>

Note: 1. Only those taxa which had at least 2% of the sequences assigned to it (in either samples) were considered for comparison
2. For each taxa, the relative ratio was obtained by dividing the percentage of a taxon observed in the malnourished sample by the percentage of the same taxon in the healthy sample.
3. Those taxa having a relative ratio of more than 1.5 or not present in the healthy sample were tagged as ‘High in malnourished’, while those taxa which had a relative ratio of less than 0.6 were tagged as ‘High in healthy’.
(IcmF, ImpB, ImpC, ImpG, ImpH and ImpJ) of Type VI secretion system. IcmF component is similar to VasK, a protein known to play a key role in cell surface recognition and adherence of bacterial pathogens to host cells [16,17]. ImpB and ImpG have been shown to be similar to *V. cholerae* secretion systems components, namely VCA0107 and VCA0111 respectively.

The analysis of malnourished specific PEGs classified under the various subsystem categories reveals a comprehensive set of PEGs belonging to interlinked pathways or subsystems that may work together and play a critical role in contributing to the pathogenicity of *Campylobacter* species in the gut of the malnourished child.

Table 2 Comparison of the taxonomic assignments obtained (using SPHINX algorithm) for sequences specific to malnourished and healthy samples (at the taxonomic level of order)

<table>
<thead>
<tr>
<th>Order Name</th>
<th>% of sequences assigned</th>
<th>Malnourished sample (X)</th>
<th>Healthy Sample (Y)</th>
<th>Relative Ratio (X/Y)</th>
<th>Inference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Campylobacterales</td>
<td>78.87</td>
<td>0.06</td>
<td></td>
<td>1235.57</td>
<td>High in malnourished</td>
</tr>
<tr>
<td>Clostridiales</td>
<td>11.4</td>
<td>0.1</td>
<td></td>
<td>119.1</td>
<td>High in malnourished</td>
</tr>
<tr>
<td>Lactobacillales</td>
<td>0.79</td>
<td>3.19</td>
<td></td>
<td>0.25</td>
<td>High in healthy</td>
</tr>
<tr>
<td>Enterobacteriales</td>
<td>0.11</td>
<td>5.07</td>
<td></td>
<td>0.02</td>
<td>High in healthy</td>
</tr>
<tr>
<td>Pseudomonadales</td>
<td>0.04</td>
<td>5.27</td>
<td></td>
<td>0.01</td>
<td>High in healthy</td>
</tr>
<tr>
<td>Chloroflexi</td>
<td>0.02</td>
<td>4.44</td>
<td></td>
<td>0</td>
<td>High in healthy</td>
</tr>
<tr>
<td>Desulfovibrionales</td>
<td>0</td>
<td>4.44</td>
<td></td>
<td>0</td>
<td>High in healthy</td>
</tr>
<tr>
<td>Bifidobacteriales</td>
<td>0</td>
<td>6.98</td>
<td></td>
<td>NA</td>
<td>High in healthy</td>
</tr>
<tr>
<td>Xanthomonadales</td>
<td>0</td>
<td>3.21</td>
<td></td>
<td>NA</td>
<td>High in healthy</td>
</tr>
<tr>
<td>Planomycetaceles</td>
<td>0</td>
<td>2.13</td>
<td></td>
<td>NA</td>
<td>High in healthy</td>
</tr>
<tr>
<td>Halobacteriales</td>
<td>0</td>
<td>12.24</td>
<td></td>
<td>NA</td>
<td>High in healthy</td>
</tr>
<tr>
<td>Actinomycetaceles</td>
<td>0</td>
<td>19.72</td>
<td></td>
<td>NA</td>
<td>High in healthy</td>
</tr>
<tr>
<td>Burkholderiales</td>
<td>0</td>
<td>19.39</td>
<td></td>
<td>NA</td>
<td>High in healthy</td>
</tr>
<tr>
<td>Rhizobiales</td>
<td>0</td>
<td>2.39</td>
<td></td>
<td>NA</td>
<td>High in healthy</td>
</tr>
</tbody>
</table>
Conclusion

This study reveals a model of the gut microbiome of a malnourished child residing in an urban slum setting in Kolkata where children are constantly exposed to enteric pathogens because of poor sanitation and hygiene and due to consumption of contaminated drinking water. Life expectancy in such resource limited setting is intriguingly dependent on physiological mechanisms of the child to prevent the swarm of extraneous pathogens. The overall differences between microbial communities residing in the gut of the malnourished and healthy child are illustrated in Figure 4. The intestinal microflora of the malnourished child when compared to the healthy child is interpreted as aberrant gut microflora. Such an aberration leads to a sub-clinical disorder characterized by inflammation and modest malabsorption. The sequel of events following the continued aberration of the intestinal microflora includes unchecked bacterial proliferation, concurrent infection, disruption in community dynamics of commensal intestinal flora and impaired immunity. Each of these events would exacerbate the other. The further reduction in diversity of essential bacterial phylotypes in the gut would result in declining gut function and competition for nutrients resulting in the downward spiraling of the child’s health. Maintenance of a healthy gut microflora as observed in the healthy child included in this study therefore seems to ensure the ability to exclude pathogens and perform its normal functioning. Apart from the current focus on dietary solutions, prevention, treatment and management of malnutrition should aim to protect the normal gut flora from infection by enteric pathogens and allochthonous microflora. An interesting question is ‘Despite living under similar conditions of hygiene and sanitation, why does a child become malnourished while the other remains normal?’. To completely understand the role of gut microbiota in malnutrition in millions of children across the globe, other models from different socioeconomic backgrounds, geographically distinct locations and diverse age groups need to be investigated.

Methods

Samples and DNA extraction

Faecal samples were collected aseptically in sterile stool containers from a healthy and a severely malnourished infant. Both the infants were 16 month old females. Samples were transported to the laboratory using frozen ice packs and were immediately stored at -80°C in 200 mg aliquots until further processing. It was ensured that no antibiotics were administered to the children for at least 3 months prior to sample collection. DNA was extracted using the bead beating method and subsequently using the Qiagen Stool DNA Mini kit. Qiagen stool lysis buffer was added to the frozen samples and used for whole community DNA extraction. 300 mg of 0.1 mm zirconia/silica beads (BioSpec Products) were added to each tube and microbial cells were then lysed by mechanical disruption with a bead beater (BioSpec Products) set on high for 2 mins. The DNA was precipitated using ethanol after removal of the inhibitors using the Qiagen Inhibi-TEX tablets and treatment with proteinase K. DNA was

Table 3 Functional analysis of sample (malnourished/healthy) specific sequences

<table>
<thead>
<tr>
<th>Subsystem category</th>
<th>% of Malnourished specific sequences (X)</th>
<th>% of Healthy specific sequences (Y)</th>
<th>Relative Ratio (X/Y)</th>
<th>Relative Ratio (Y/X)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motility and Chemotaxis</td>
<td>4.18</td>
<td>0.04</td>
<td>104.5</td>
<td>0.01</td>
</tr>
<tr>
<td>Membrane Transport</td>
<td>2.12</td>
<td>0.21</td>
<td>10.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Respiration</td>
<td>5.37</td>
<td>1.28</td>
<td>4.2</td>
<td>0.24</td>
</tr>
<tr>
<td>Virulence</td>
<td>9.49</td>
<td>5.28</td>
<td>1.8</td>
<td>0.56</td>
</tr>
<tr>
<td>Stress Response</td>
<td>1.95</td>
<td>1.83</td>
<td>1.07</td>
<td>0.94</td>
</tr>
<tr>
<td>Protein Metabolism</td>
<td>12.1</td>
<td>11.44</td>
<td>1.06</td>
<td>0.95</td>
</tr>
<tr>
<td>Amino Acids and Derivatives</td>
<td>8.45</td>
<td>8.54</td>
<td>0.99</td>
<td>1.01</td>
</tr>
<tr>
<td>DNA Metabolism</td>
<td>7.01</td>
<td>7.37</td>
<td>0.95</td>
<td>1.05</td>
</tr>
<tr>
<td>Cofactors, Vitamins, Prosthetic groups and pigments</td>
<td>7.07</td>
<td>7.71</td>
<td>0.92</td>
<td>1.09</td>
</tr>
<tr>
<td>Cell wall and Capsule</td>
<td>6.52</td>
<td>7.38</td>
<td>0.88</td>
<td>1.13</td>
</tr>
<tr>
<td>Carbohydrates</td>
<td>5.78</td>
<td>7.95</td>
<td>0.73</td>
<td>1.38</td>
</tr>
<tr>
<td>Clustering-based Subsystems</td>
<td>11.5</td>
<td>16.24</td>
<td>0.71</td>
<td>1.41</td>
</tr>
<tr>
<td>RNA Metabolism</td>
<td>5.11</td>
<td>7.58</td>
<td>0.67</td>
<td>1.48</td>
</tr>
<tr>
<td>Cell Division and Cell Cycle</td>
<td>1.61</td>
<td>2.43</td>
<td>0.66</td>
<td>1.51</td>
</tr>
<tr>
<td>Unclassified</td>
<td>2.26</td>
<td>4.45</td>
<td>0.51</td>
<td>1.97</td>
</tr>
<tr>
<td>Nucleosides and Nucleotides</td>
<td>3.38</td>
<td>6.66</td>
<td>0.51</td>
<td>1.97</td>
</tr>
</tbody>
</table>

Note: Only those subsystem categories which had at least 1.5% of the sequences assigned to it were considered for comparison.
purified by binding on a QIAamp mini spin column and subsequent washing with wash buffers and finally eluted in 200 μl TE (pH 8.0). The quantity and quality of purified DNA was assessed spectrophotometrically and also by using agarose gel electrophoresis.

Community Metagenome Sequencing
Total community DNA extracted from each stool specimen was directly sequenced on individual pyrosequencing slides on a ROCHE 454 GS FLX (Roche Diagnostics, Inc. Basel, Switzerland) sequencing instrument. While the healthy child sample was sequenced using the GS FLX chemistry, the malnourished child sample was sequenced using the Titanium chemistry. The distinct difference in read lengths in these two methods did not affect the downstream bioinformatic processes as we undertook an assembly independent analysis of the two metagenomes.

Bioinformatic analysis
Low quality and short sequences were first removed from both data sets using in-house scripts. Duplicated
sequences, a known artifact in pyrosequencing data, were also removed. Using Eu-Detect algorithm (http://metagenomics.atc.tcs.com/Eu-Detect), sequences of probable eukaryotic origin were identified. Sequences originating from human DNA were then identified by performing a BLASTn search of these sequences against human genome sequences. Sequences having at least 80% identity (spread over at least 80% of the length) to human sequences were removed from subsequent analyses.

All reads in both data sets were taxonomically classified using the SPHINX algorithm [5]. Analysis of these assignments at various taxonomic levels was performed by first collapsing all assignments at a desired taxonomic level, and subsequently enumerating the number of assignments to various taxa at that level.

Sequences specific to either the malnourished or the healthy child data set were identified using the HabiSign algorithm (TSG et al., manuscript communicated). This algorithm identifies sequences specific to a data set by first mapping all sequences to pre-computed points in feature vector space. Subsequently, sequences specific to a data set are identified by finding regions in the same feature vector space that are observed to be selectively over-mapped by sequences belonging to that data set. For each data set, the taxonomic affiliations of the identified specific sequences were inferred from the results of SPHINX analysis.

Functional categorization of the sequences identified as specific to each data set were performed by submitting these sequences to MG-RAST server (http://metagenomics.nmpdr.org/). The SEED platform hosted on this server contains all protein sequences classified under various sub-systems [12]. The specific sequences were compared against the proteins in various sub-systems using BLASTx with an e-value cutoff of e-10, percent identity greater than 66% and an alignment length of greater than 50 bases. The percentage of specific sequences tagged to various functional categories were obtained and subsequently compared.

Consent
Written informed consent was obtained from the parents of the two children for publication of this case.
report and accompanying images. A copy of the written consent is available for review by the Editor-in-Chief of this journal.

Abbreviations
MDG-1: First Millennium Development Goal; MG-RAST: Metagenome Rapid Annotation using Subsystems Technology; PEGs: Protein Encoding Genes

Acknowledgements
We hereby acknowledge the efforts, cooperation and support extended by Roche diagnostics India Pvt. Ltd for undertaking the sequencing of the metagenomic samples. We thank Varun Mehra, Sudha Chadraram and Dinakar Komanduri for their help during the computational analysis.

Authors’ contributions
SSG prepared DNA from the faecal samples and collated the information on the sequences, SK collected the samples and information for categorizing the child as malnourished, MMH, TSG and SSM performed detailed bioinformatics analysis. SSG, MHM, TSG, GBN and SSM analyzed the results and wrote the paper. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Received: 21 April 2011 Accepted: 20 May 2011 Published: 20 May 2011

References