**Journal of Genetics** (2021)100:84
https://doi.org/10.1007/s12041-021-01326-x



**RESEARCH NOTE**

# Gut microbiome architecture of wild greater one-horned rhinoceros: a vulnerable species from Kaziranga National Park, India

PARIKSHIT KAKATI[1], SUMAN KALYAN PAINE[2]*, CHANDRA KANTA BHATTACHARJEE[3], CHANDRIKA BHATTACHARYYA[2], AMIT SHARMA[1], DEBABRATA PHUKAN[4], NAGENDRA NATH BARMAN[5] and ANALABHA BASU[2]*

[1]*WWF-India, A-16/103, Game Village, Basistha, Brahmaputra Landscape, Wildlife and Habitat Division, Guwahati, India*
[2]*National Institute of BioMedical Genomics, Kalyani 741 251, India*
[3]*ICMR-Regional Medical Research Centre, NE Region, Dibrugarh 786 010, India*
[4]*Kaziranga National Park, Assam Forest Department, Govt. of Assam, Bokakhat, India*
[5]*College of Veterinary Science, AAU, Khanapara, Guwahati 781 022, India*
*For correspondence. E-mail: Analabha Basu, ab1@nibmg.ac.in; Suman Kalyan Paine, painesuman01@gmail.com.

Received 28 October 2020; revised 7 July 2021; accepted 11 July 2021

**Abstract.** *Rhinoceros unicornis*, also known as the greater one-horned rhinoceros (GoHR), is a vulnerable wildlife species found in the Indian subcontinent with an estimated global population of 3582, of which an estimated 2995 resides in India. The Kaziranga National Park of Assam is the home to ∼80.56% of the GoH population in India. Recent advances in genetics and microbial studies underscored the importance of gut microbial symbiosis as a crucial factor for host metabolic health and environmental interaction, particularly for higher mammals. Alteration of the normal microbiome can also be an indicator of chronic disease and infection. Freshly voided dung samples from nine dung heaps of free ranging or wild *GoH* rhinoceros were collected from Kaziranga National Park for mapping the gut microbial architecture through 16S-metagenomic approach. In our sample, the *GoH* gut harbours 168.8±12.55 (SE) bacteria-specific OTUs belonging to 21 phyla of which the gram-negative Proteobacteria is the most abundant phyla. Other abundant phylas found in the GoH gut are Firmicutes and Bacteroidetes. Although the GoH rhinoceros gut can utilize fibrous plant by microbial fermentation, the aerobic, nonfermenting *Acinetobacter* (20.7%), *Stenotrophomonas* (17.8%) and *Brevundimonas* (9.1%) constitute about 50% of all identified genus. Functional prediction of the GoH microbiome reveals that >50% of the bacteria present are involved in metabolism followed by cellular processes and information processing. A significant proportion (>1%) are associated with different diseases. In summary, our study characterized bacterial communities of nine wild GoH to identify some unique features and its implication in disease and survival of GoH.

**Keywords.** microbiome; Kaziranga; vulnerable species and herbivore; *Rhinoceros unicornis*.

## Introduction

The greater one-horned rhinoceros (GoHR), *Rhinoceros unicornis*, is a vulnerable wildlife species found in the Indian subcontinent (Talukdar *et al*. 2008, IUCN/SSC Guidelines v. 1.0 2013). The GoHR have an estimated population of around 3747 individuals, spread across the foothills and grasslands of eastern Himalayas and Brahmaputra valley (The Rhino Research Centre, Cambridge, United Kingdom 2019, http://www.rhinoresourcecenter.com; WWF Report 2017, https://www.worldwildlife.org/species/greater-one-horned-rhino). Among the Indian population of about 2995 GoHR (80.00% of the global population), around 2664 of them are found in Assam and of these around 80.56% (2413 as per 2018 estimation) are found in Kaziranga National Park alone (The Hindu, 30 March 2018, https://www.thehindu.com/news/national/other-states/rhino-census-2018-

Parikshit Kakati and Suman Kalyan Paine contributed equally to this work.

*Supplementary Information: The online version contains supplementary material available at https://doi.org/10.1007/s12041-021-01326-x.*

Published online: 12 November 2021

kaziranga-now-has-2413-rhinos/article23393316.ece). The Kaziranga National Park of Assam occupies the centre stage to all conservation efforts for GoH, as it is the home to ~91% of the GoH population in India. Because of the vulnerable status of the species, an understanding of the pathophysiology and identification of a congenial environment for the survival of the species is of strategic importance. Recent advances in genetics and microbial studies underscored the importance of gut microbial symbiosis as a crucial factor for host metabolic health and environmental interaction, particularly for higher mammals.

The GoHR is a 'simple stomach mega herbivore' weighing up to 2.5 tones and after the elephant, it is the largest extant mammalian herbivore. It has the ability to utilize fibrous plant matter through microbial fermentation in the hindgut (Clauss *et al*. 2005). To digest the huge quantity of grass under an anaerobic environment, a specialized microbial community is essential in these herbivores (Flint 1997). The establishment of a specialized microbiome corresponding to a hindgut fermenter species in the wild is a coevolutionary process in herbivores and is dependent on the diet and habitat (Ley *et al*. 2008; Gibson *et al*. 2019). To understand the nutritional and digestive physiology as well as the functional and biological significance of the gut microbiome of GOH, we have predicted the functional profile of the microbial communities using 16s rRNA marker gene sequences. Our analysis reveals that the majority (>50%) of the bacteria in the microbiome are related to metabolism. Apart from that, we found around 18% of the gut bacteria are functionally predicted as involved in genetic information processing, followed by cellular processes (~10%) and environmental information processing (~10%). A significant percentage (>1%) of the bacteria are predicted to be harmful and are associated with different diseases. The gut microbiome needs to be studied for the vulnerable wildlife species like GoHR. Hence, a better understanding is necessary to facilitate the planning and implementation measures to address rehabilitation, disease outbreak and the survival of the vulnerable species like the GoHR.

To characterize the complex microbial populations in the GoHR gut, we harnessed the power of high throughput massively parallel sequencing. Our present study comprehensively mapped the gut microbial architecture among wild population of GoHR from India using the 16S metagenomic approach. To the best of our knowledge, the present study is the first of its kind, which uses high throughput sequencing to characterize the complex gut-microbial populations of the GoHR from the wild natural habitat.

## Material and methods

Freshly voided dung samples (minimum 50 gm) from nine dung heaps (specifically inner masses of the dung heap to avoid possible environmental contaminants) of free ranging or wild GoHR were collected during the month of July 2018 from Kaziranga National Park. The study was approved by the Forest Department, Govt. of Assam for collecting the wild free ranging GoHRs' dung in noninvasive procedure. Dung samples (minimum 50 gm) were collected from the middle of the bolus and kept in sterile plastic containers without any preservatives, and was immediately despatched to the laboratory in an icebox. All possible efforts were made to collect freshly voided dung samples not older than the previous night. The selection was purely based on expert's inspection on the physical and visual parameters of the dung (moisture content, shine and presence of mucous layer, presence of maggots, dung beetles and external/surface fungal growth like mushrooms and toadstool). The dried samples with external fungal growth over the surface, dung heaps scratched and dispersed by wild fowls, birds and wild boars were not picked up for the study. DNA was extracted using the Qiagen Stool DNA kit. Amplicons were generated through 16S universal primer for variable regions between 3 and 4, and sequenced on Illumina-HiSeq and analysed through QIIME (v. 1.9.0) (Caporaso *et al*. 2010). Details are presented in a–c in electronic supplementary at http://www.ias.ac.in/jgenet/. Ethical approval for this study was obtained from Forest Department, Govt. of Assam and Institutional

**Table 1.** Represents the read count and QC for each sample (nine *R. unicornis*) of pair end 16S amplicon sequencing on Illumina HiSeq2500.

| Sample ID | Sequence reads of each for R1 and R2 reads | Combined pairs | Dereplicated pairs | Nonchimeric pairs | R1 (% Q > 30) | R2 (% Q > 30) | Mean read length (bp) |
|---|---|---|---|---|---|---|---|
| 10R | 1117782 | 771661 | 749837 | 539411 | 85.53 | 76.26 | 251 |
| 1R | 1330294 | 785547 | 651307 | 521009 | 85.82 | 74.64 | 251 |
| 2R | 1081725 | 754076 | 717876 | 557801 | 86.21 | 78.01 | 251 |
| 3R | 1165296 | 767913 | 742165 | 579629 | 86.45 | 77.8 | 251 |
| 4R | 918327 | 613987 | 595758 | 419410 | 85.18 | 75.11 | 251 |
| 6R | 1131494 | 801627 | 754089 | 525982 | 85.80 | 77.02 | 251 |
| 7R | 986196 | 703975 | 669620 | 473580 | 86.20 | 76.78 | 251 |
| 8R | 849181 | 525053 | 503198 | 395778 | 85.60 | 75.92 | 251 |
| 9R | 894630 | 569181 | 551190 | 441477 | 86.43 | 76.66 | 251 |

ethic committee of College of Veterinary Science, AAU, Khanapara.

## Results

In the present study, involving data from nine GoHR, we estimated that on an average, GoHR gut harbours 168.8±12.55 (SE) (range: 108–236) bacteria-specific OTUs. Sequence reads generated for the samples ranged from 1,330,294 to 894,630 (table 1) and across all samples, a total of 1,226,013 sequence reads were identified as bacterial, covering 21 phyla (table 1 in electronic supplementary material). The data reveals that overall the Proteobacteria is the most abundant phyla (minimum 19.09% and maximum 94.09%) of the individual GoHR guts, followed by Firmicutes (minimum 1.3% and maximum 60.8%) and Bacteriodets (minimum 1.2% and maximum 19.6%) among the samples (figure 1).

The bacteria-specific OTUs and alpha diversity indices (Shannon–Weaver, Simpson, Inverse Simson, Pielou's evenness (J) and Fisher alpha) for each individual GoHR is presented in table 2. The alpha diversity of a calf (7R) is different from the adult GoHR and it exhibits the lowest alpha diversity index. We computed the Bray–Curtis β-diversity index (BCDI) among the nine GoHR that are presented in figure 2. The BCDI among the nine wild GoHR ranged from 0.38 to 0.96.

**Table 2.** Alpha diversity index of *R. unicornis.*

| | OTUs | Shannon | Simpson | Fisher alpha | Inverse Simpson |
|---|---|---|---|---|---|
| 1R | 197 | 1.30 | 0.54 | 14.44 | 2.10 |
| 2R | 108 | 1.40 | 0.75 | 13.90 | 4.13 |
| 3R | 161 | 2.30 | 0.79 | 14.05 | 4.86 |
| 4R | 149 | 2.50 | 0.87 | 9.89 | 7.85 |
| 5R | 182 | 2.60 | 0.89 | 11.32 | 9.13 |
| 6R | 135 | 2.70 | 0.85 | 12.01 | 6.70 |
| 7R | 236 | 0.90 | 0.41 | 7.72 | 1.69 |
| 8R | 191 | 1.60 | 0.67 | 11.23 | 3.09 |
| 9R | 160 | 3.09 | 0.90 | 17.60 | 10.05 |

A total of 39 OTUs, homogeneous across all nine GoHR, covering 90.4% (1,108,929 OTU reads) of total 1,226,013 bacterial sequence reads among all nine samples have been treated as the core OTUs (figure 3). Among the 39 core OTUs for nine GoHR, 18 OTUs identified the genus that covers 925,621 sequence reads (83.4%). Within the core OTUs, the genus *Acinetobacter* (20.7%), *Stenotrophomonas*

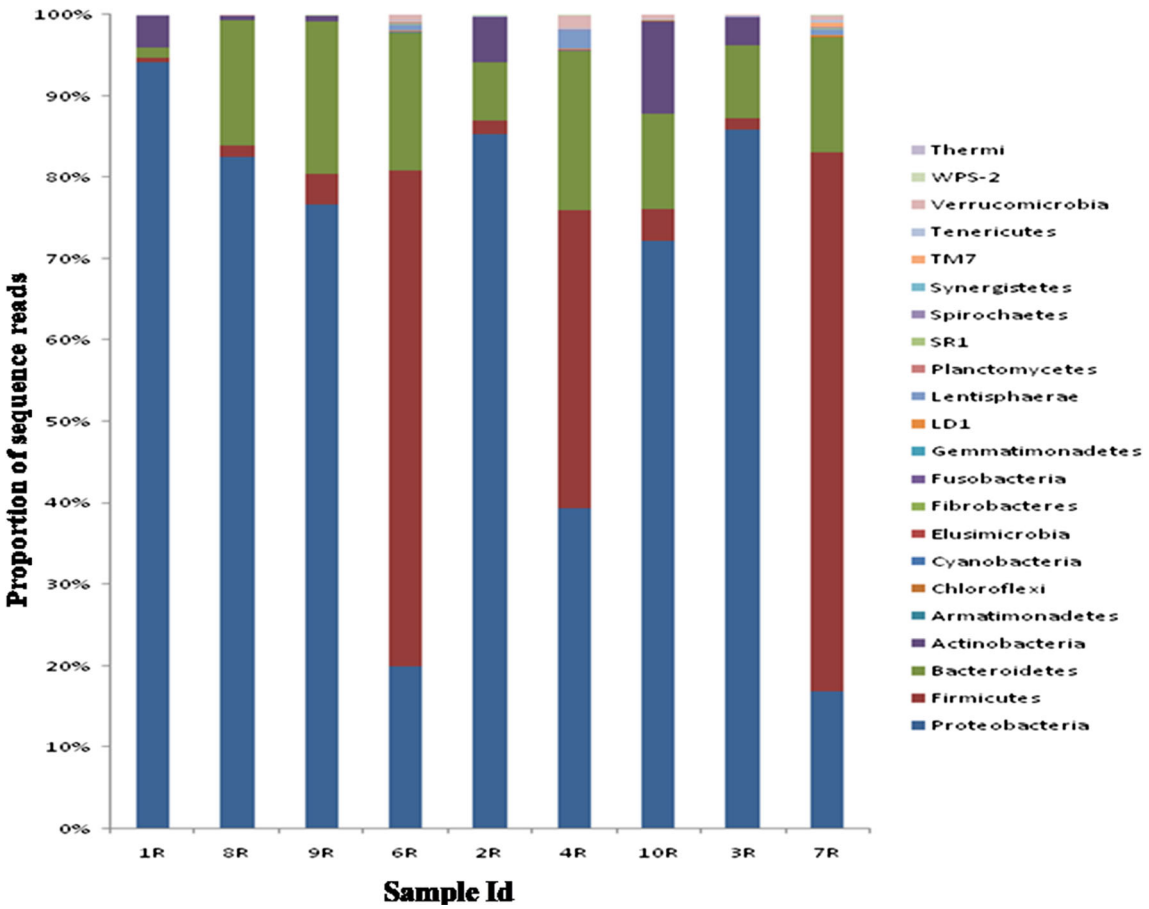


**Figure 1.** Represents the abundance of 21 annotated phyla for each nine wild *R. unicornis.*

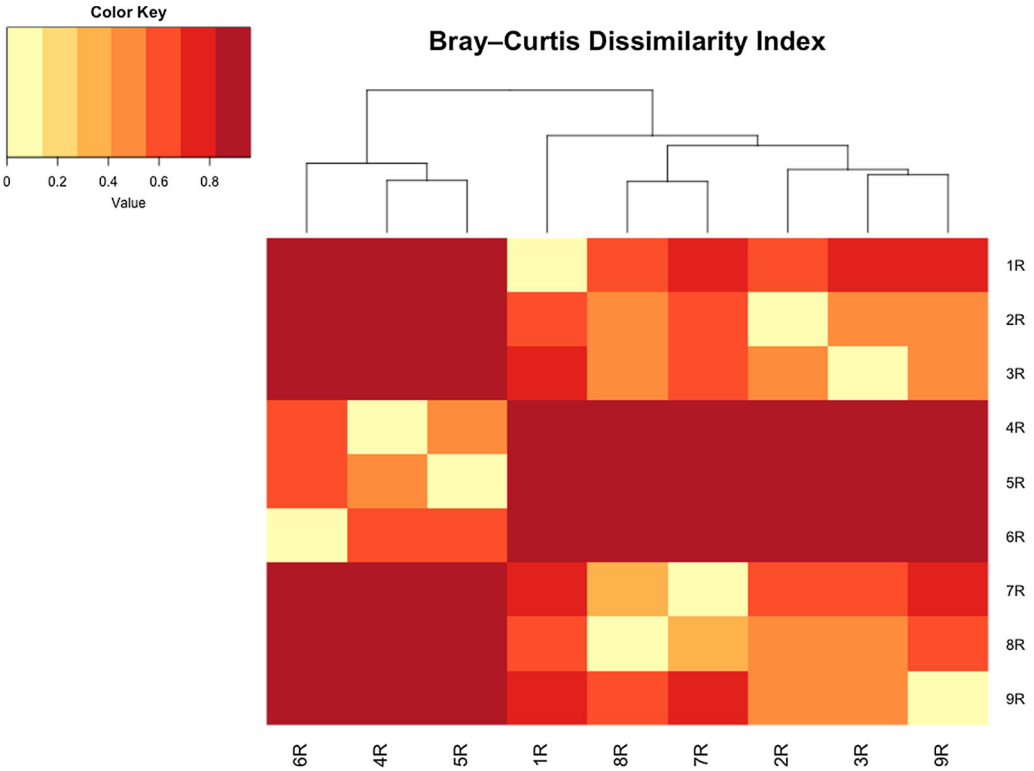


**Figure 2.** BCDI among the nine GoHR.

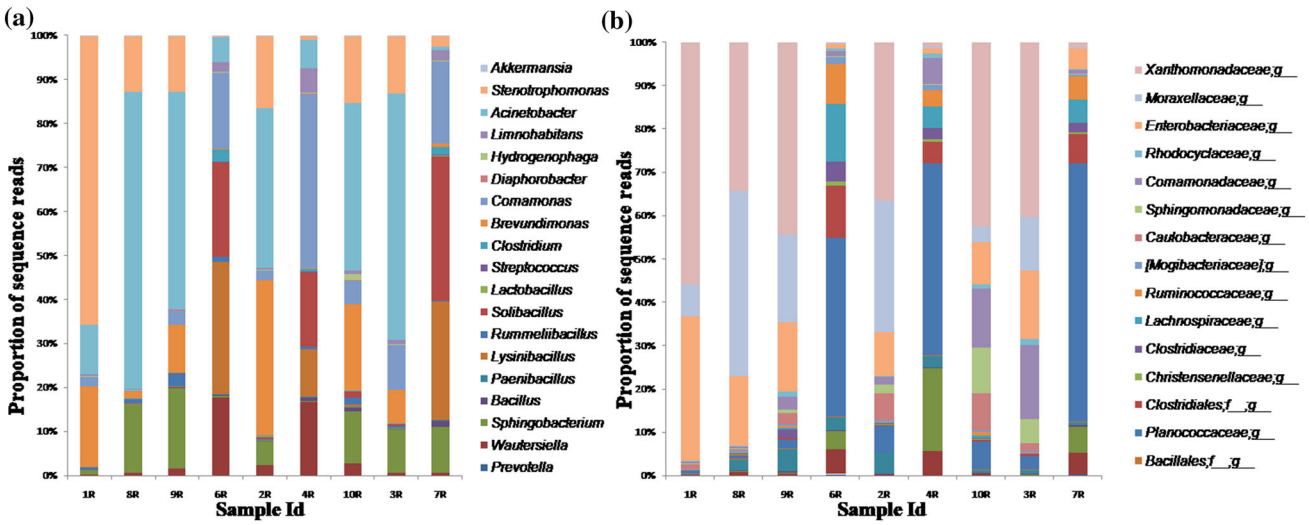


**Figure 3.** Thirty-nine OTUs are homogenus across the all nine *R. unicornis* that covers 90.4% of total 1,226,013 sequence reads that assigned for bacterial sequence among all nine samples, may exhibit as core. (a) OTUs identified as bacterial genus that are present in all nine *R. unicornis*. Among the 39 core OTUs, 18 OTUs identified the genus that covers 83.4% sequence reads. (b) OTUs identified as bacterial sequence as unidentified genus that are present in all nine *R. unicornis.* Among the 39 core OTUs, 16.5% bacterial sequence specific OTUs were unable to identify the genus.

(17.8%), *Brevundimonas* (9.1%) *Comamonas* (6.3%), *Solibacillus* (5.3%), *Lysinibacillus* (5.3%), *Sphingobacterium* (5.1%), *Wautersiella* (3.08%) are present in abundant manner, whereas *Limnohabitans*, *Rummeliibacillus*, *Clostridium*, *Bacillus*, *Hydrogenophaga*, *Lactobacillus*, *Streptococcus*, *Paenibacillus*, *Diaphorobacter*, *Prevotella*, *Akkermansia* also coexist as core with <1% among the all nine *R. unicornis* (figure 3a). Apart from the 18 genus-specific OTUs, we found another set of 21 bacteria-specific OTUs constituting 183,308 sequence reads (16.5% of total read). These OTUs lacked the resolution to identify the genus of the bacteria, but belonged to different taxonomic

groups like Coriobacteriaceae, Flavobacteriaceae, Planococcaceae, Xanthomonadaceae, Comamonadaceae, Enterobacteriaceae, Rhodocyclaceae, Moraxellaceae, Mogibacteriaceae, Christensenellaceae and Bacteroidales (figure 3b). For the nine GoHR, the number of OTUs that were identified as core, but lacked the resolution to identify the genus ranged from minimum 124 (0.0001%) sequence reads to a maximum of 53,088 (4.3% of total read) sequence reads. The OTUs belonging to the families Planococcaceae (4.3% of total read), Xanthomonadaceae (3.2% of total read) and Moraxellaceae (1.7% of total read) are the major ones for which the genus was not identifiable. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences through phylogenetic investigation of communities by reconstruction of unobserved states (PICRUSt) reveals that the majority (50.2 ± 1.56%) of the bacteria in the microbiome are related to metabolism. Apart from that, we found around 18 per cent (17.8± 0.06) of the gut bacteria are functionally predicted as involved in genetic information processing. Other important functions predicted are cellular processes (9.7 ± 0.07%) and environmental information processing (9.9 ± 0.01%). A significant percentage (>1%) of the bacteria are predicted to be harmful and are associated with different diseases (figure 4).

## Discussion

It is essential to initiate the studies for understanding the mutual symbiotic gut bacterial association and its dysbiosis through diseases and pathogens affecting the GoHR as a considerable number of cases of GoHR deaths due to pathogenic infection go unaddressed because of our ignorance. As per the IUCN's Disease and Parasite consideration (5.1.6), surveillance of source populations can establish the potential pathogen community present among the individuals of the population, thus enhancing our understanding of host–pathogen interaction for the species. Both of the above warrants a comprehensive characterization of wild GoHR gut microbiome. This study documents that the *R. unicornis* gut is dominated by gram-negative Proteobacteria, Firmicutes and Bacteriodetes. In our study, 0.14% of the GoH gut microbiota remains unclassified. The GoH rhino gut microbiome potentially differs from the white and black rhino, whose gut microbiome is dominated by Firmicutes and Cyanobacteria (Bian *et al*. 2013; Gibson *et al*. 2019). However, Proteobacteria dominated herbivore gut has been documented in extinct mega herbivores like mammoth (76.2%) and Wooly rhinoceros (19.8%) (Gibson *et al*. 2019). In these extinct species, Proteobacteria was followed by Firmicutes and Actinobacteri (Talukdar *et al*. 2008). In domesticated herbivores, a distinct microbial architecture has been observed. For instance, the gut of dairy cattle is dominated by Firmicutes (70.1%) followed by Bacteroidetes (8.1%), Actinobacteria (7.3%) and Proteobacteria (2.5%). Surprisingly though, dietary alteration including antibiotics did not reveal significant alteration of GI tract microbial architecture. 'Firmicutes dominated gut microbial architecture' was also documented in several carnivorous animals like leopard cats (63%), otters (88%) and raccoon dogs (90%), while Proteobacteria is present in high abundance among

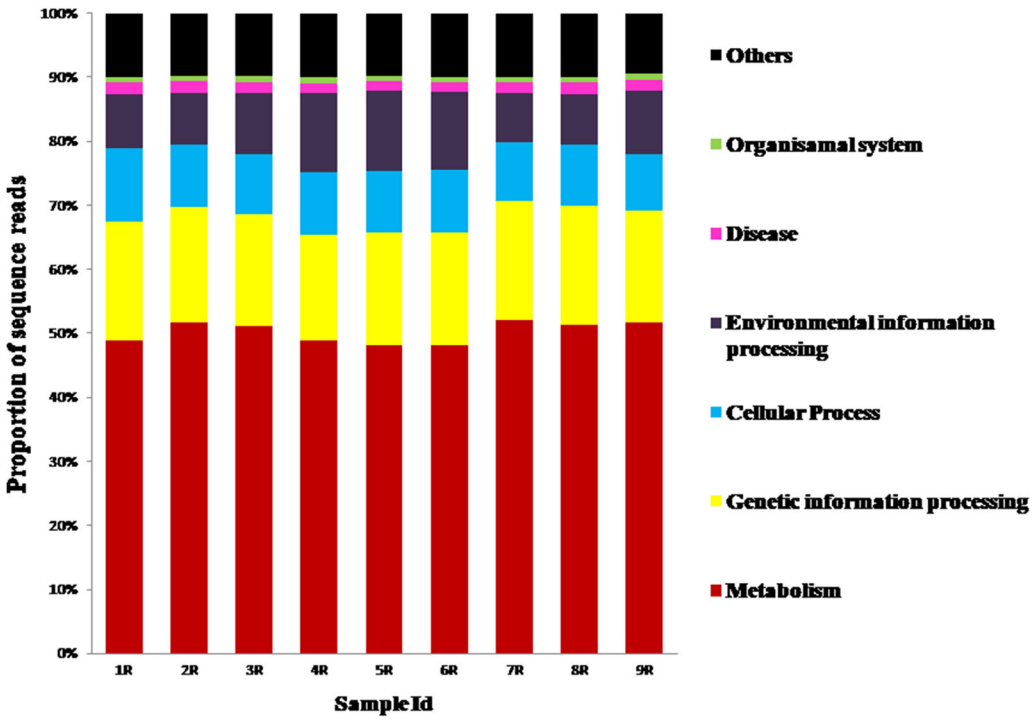


**Figure 4.** Represents the predictive functional profiling of microbial communities through PICRUSt.

leopard cats (range 10–80%). Habitat specific difference in bacterial community has also been documented for animals; e.g., the wild red panda gut is Proteobacteria dominated (40–80%) while in captivity it is dominated by Firmicutes (80–90%). The GoHR gut-microbiome also harbours soil bacteria and bacteria that are present in plants like *Solibacillus* (5.3%), *Lysinibacillus* (5.3%), *Sphingobacterium* (5.1%) *Limnohabitans*, *Rummeliibacillus* and *Hydrogenophaga* that are associated with plant are also present in <1% abundance. The presence of soil and plant associated bacteria in GoHR stool possibly exhibits its poor digestive herbivore nature. These bacteria are hence likely to be habitat specific, like in this case all the samples were collected from grasslands of Brahmaputra valley. *Brevundimonas*, *Prevotella* and *Clostridia* that are present in GoHR gut are associated with plant-derived fibre degradation among the herbivores (Mardanov *et al*. 2012). *Lactobacillus* and *Streptococcus* that are abundant inhuman gut was also present in limited abundance in all nine GoHR samples. These bacteria are likely to be the core and are crucial for energy metabolism and growth.

In summary, the study documents the composition of bacterial communities in the faeces of nine wild *R. unicornis* from the grasslands of Brahmaputra valley. Present data reveals the presence of a unique bacterial population that comprises admixture of plant, soil and animal associated bacteria. The core gut bacteria of *R. unicornis* exhibit its own distinctive microbial architecture that comprehensively differ from other rhinos (black rhino and white rhino) as well as other herbivores like horses, swine and dairy cattle (Flint *et al*. 2008; Daly *et al*. 2012; Mardanov *et al*. 2012). Although there are differences in the proportion of different bacteria in the GoHR, it is to be noted that after the functional prediction of the gut microbiome, the individual variation gets considerably reduced (figure 4). The data reveals that about 50% of sequence read corresponds to metabolism, 20% reads are involved to explain genetic information processing and 10% of reads are carrying the information on cellular process, environmental information processing.

This study initiates investigation of *R. unicornis* gut microbiome using a high throughput 16S metagenomic sequencing to map its gut. The observations may be crucial to understand the bacterial ecosystem of this vulnerable wildlife species and would enable us to better understand the involvement of those bacteria in varied range of functional prediction that includes host metabolism, genetic information processing, cellular network and diseases (figure 4). A comprehensive catalogue of the gut microbiome of GoHR provides an opportunity for disease surveillance using the easy and noninvasive route of gut microbiome analysis. It also initiates the required groundwork for possible habitat identification and rehabilitation of the species. The datasets generated for the current study are available with the corresponding author.

**Acknowledgements**

We would like to acknowledge the Chief Wildlife Warden of Assam for allowing us to initiate the study. We are thankful for the help and needful support provided by the Field Director, Kaziranga National Park, Divisional Forest Officer, East Assam Wildlife Division, Range Forest Officers and the forest guards for providing necessary assistance for carrying out this study and collect necessary samples. The authors sincerely acknowledge Mrs Soumita Dutta for laboratory assistance. We would like to thank Sri Ravi Singh, SG & CEO WWF India for his constant motivation and support to undertake challenging works. We thank the team members of WWF India especially Dr Dipankar Ghose and Dr Anupam Sarmah for their help and support during the study. This study was possible because of the support of our donors, especially WWF-US. The study was funded by World Wildlife Fund (WWF) US grant number PRO-927.

**Author contributions**

PK, SKP and AB were involved in the study design and data generation, and analysis. DP, AS and PK conducted the field work and sampling. NNB facilitated sample processing and coordinating between the two laboratories of College of Veterinary Science, AAU, Khanapara and Regional Medical Research Centre. CKB and CB were involved in data generation and analysis. All authors were actively involved in manuscript preparation.

## References

Bian G., Ma L., Su Y. and Zhu W. 2013 The microbial community in the feces of the white rhinoceros (Ceratotheriumsimum) as determined by barcoded pyrosequencing analysis. *PLoS One* **8**.

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. Methods* **7**, 335.

Clauss M., Polster C., Kienzle E., Wiesner H., Baumgartner K., Houwald Von *et al*. 2005 Studies on digestive physiology and feed digestibilities in captive Indian rhinoceros (Rhinoceros unicornis). *J. Anim. Physiol. Anim. Nutr.* **89**, 229–237.

Daly K., Proudman C. J., Duncan S. H., Flint H. J., Dyer J. and Shirazi-Beechey S. P. 2012 Alterations in microbiota and fermentation products in equine large intestine in response to dietary variation and intestinal disease. *British J. Nutr.* **107**, 989–995.

Flint H. J. 1997 The rumen microbial ecosystem—some recent developments. *Trends Microbiol.* **5**, 483–488.

Flint H. J., Bayer E. A., Rincon M. T., Lamed R. and White B. A. 2008 Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. *Nat. Rev. Microbiol.* **6**, 121.

Gibson K. M., Nguyen B. N., Neumann L. M., Miller M., Buss P., Daniels S. B. *et al*. 2019 Gut microbiome differences between wild and captive black rhinoceros–implications for rhino health. *Sci. Reports* 1–11.

IUCN/SSC 2013 Guidelines for reintroductions and other conservation translocations. Version 1.0 IUCN Monographic Series 10386.

Ley R. E., Hamady M., Lozupone C., Turnbaugh P. J., Ramey R. R., Bircher J. S. *et al*. 2008 Evolution of mammals and their gut microbes. *Science* **320**, 1647–1651.

Mardanov A. V., Bulygina E. S., Nedoluzhko A. V., Kadnikov V. V., Beletskii A. V. Tsygankova S. V. *et al*. 2012 Molecular

analysis of the intestinal microbiome composition of mammoth and woolly rhinoceros. In *Biochemistry and biophysics* (ed. Doklady), vol. 445, pp. 203. Springer Science.

Talukdar B. K., Emslie R., Bist S. S., Choudhury A., Ellis S., Bonal B. S. *et al*. 2008 Rhinoceros unicornis. The IUCN Red List of Threatened Species.

Corresponding editor: T. N. C. VIDYA

**Supplementary data:**

**Gut microbiome architecture of wild greater one horned rhinoceros, a vulnerable species from Kaziranga National Park, India**

Parikshit Kakati, Suman Kalyan Paine, Chandra Kanta Bhattacharjee, Chandrika Bhattacharyya, Amit Sharma, Debabrata Phukan, Nagendra Nath Barman and Analabha Basu

**a) 16S Metagenome Sequencing**

Stool samples were homogenized with 1XPBS and incubated lysozyme solution (Sigma Aldrich) at 37°C for 1 hour. 500ul Inhibitex buffer (Qiagen) was mixed vigorously and incubated at $70^0$C and $90^0$C for 20 min. and 5 min., respectively. DNA was extracted using the Qiagen stool DNA Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions and eluted in 50 µl elution buffer provided with the kit.

**b) Gene-specific sequences**

The gene specific sequences used in this protocol target the 16S V3 and V4 region. The full lengthprimer sequences, using standard IUPAC nucleotide nomenclature, to follow the protocoltargeting this region are:

Primer sequences:
16S Forward: 5' AGAGTTTGATCCTGGCTCAG 3'
16S Reverse: 5' GGTTACCTTGTTACGACTT 3'
V3V4F 5' CCTACGGGNGGCWGCAG 3'
V3V4R 5' GACTACHVGGGTATCTAATCC 3'

Each sequenced sample is prepared according to the Illumina16S Metagenomic Sequencing Library protocols. The quantification of DNA and the DNA quality is measured by PicoGreen and Nanodrop. 2ng Genomic DNA was used in the first amplification step using the following conditions: 5pmol forward tailed target specific primer; 5pmol reverse tailed target specific primer; and Herculase II polymerase(Agilent). The PCR for each variable region was carried out in triplicate in a 25ul reaction in a thermal cycler PCR system with the following parameters: initial denaturation at 95°C for 3mins, followed by 25cycles of 95°C for 30s, 55°C for 30s, and 72°C for 30s with a final extension at 72°C for 5minand then stored at 4 °C hold. The amplicon libraries were cleaned to remove excess nucleotides, salts and enzymes using 20ul of the AgencourtAMPure XP system (Beckman Coulter Genomics) and

eluted in 25ul of TE buffer. The 10ul of the first step reaction was submitted to a second amplification step using the following conditions: Nextera XT Index Primer (N7xx); Nextera XT Index Primer (S5xx); Herculase II polymerase(Agilent). The PCR for each variable region was carried out in triplicate in a 25ul reaction in the above-mentioned thermal cycler with the following parameters: initial denaturation at 95°C for 3min, followed by 8cycles of 95°C for 30s, 55°C for 30s, and 72°C for 30s with a final extension at 72°C for 5min and then stored at 4 °C hold. The amplicon libraries were cleaned to remove excess nucleotides, salts and enzymes using 20ul of the Agencourt AMPure XP system (Beckman Coulter Genomics) and eluted in 25ul of TE buffer.The final purified product is then quantified using qPCR according to the qPCR Quantification Protocol Guide Guide (KAPA Library Quantificatoin kits for IlluminaSequecing platforms) and qualified using the Tape Station DNA screentape D1000 (Agilent). In preparation for cluster generation and sequencing, pooled libraries are denatured with NaOH, diluted with hybridization buffer, and then heat denatured before sequencing. Each run include a minimum of 5% PhiX to serve as an internal control for these low- diversity libraries. These libraries were pooled based on the data required and sequenced on HiSeq 2500 sequencer using Rapid SBS kit V2 (500 cycles) to generate 250 paired-end reads.

**c) Analysis**

The initial quality control was performed based on the FASTQC reports (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) for each sequence file. The files were then demultiplexed for the MID sequences and trimmed using the following filter criteria: Number of Ambiguous bases = 0–2, QV per base = 20–25, Minimum sequence length = 150–200, Maximum Sequence Length = 250–300 and Number of Homopolymers = 4–6 based on the nature of the data. The 16's sequence read files were analysed for the 28 samples on QIIME (Version 1.9.0). Chimera filtering was performed using UCHIME. Maximum and minimum read count among the samples were (pair end read) 31.6 million and 27.2 million, respectively with an average Q30 (phred score) of 79.8%. The Green genes database was used for Taxonomy assignment. Operational Taxonomic Units (OTUs) picking was performed using the *pick_otus.py* command with the default UCLUST algorithm. The UCLUST algorithm uses the USEARCH algorithm to assign sequences to a cluster. The USEARCH algorithm works by searching a query sequence against target sequences and recording the k-mers in common between the two sequences. Rather than inferring sequence similarity as the number of matching k-mers between a query and target sequence, USEARCH arranges the target sequences in decreasing order of the number of unique k-mers shared between the two sequences. The query sequences

are arranged into clusters. Each cluster centroid shares a level of similarity below a set identity threshold level with each other centroid. The remaining query sequences are then assigned to a centroid (target sequence) based on identity threshold using the USEARCH algorithm. If the query sequence does not share similarity with a centroid above the threshold a new cluster is created. The most abundant read in each OTU was selected as the representative sequence; this step was performed using *pick_rep_set.py. Assign_taxonomy.py* was used for the classification of each of the representative sequences. All the statistical analysis and heatmaps were made by R packages [R version 3.2.3 (2015-12-10) Copyright (C) The R Foundation for Statistical Computing].

**Table 1.** Represents the read count and abundance of 21 annotated Phyla among wild nine

*Rhinoceros unicornis* of pair end 16S amplicon Sequencing on Illumina HiSeq2500.

| Phyla | Minimum | Maximum | Mean | Std. deviation | Std. error | Sum |
|---|---|---|---|---|---|---|
| Acidobacteria | 0 | 28 | 3.333 | 9.274 | 3.091 | 30 |
| Actinobacteria | 24 | 11956 | 3815 | 4624 | 1541 | 34339 |
| Armatimonadetes | 0 | 28 | 6.222 | 9.821 | 3.274 | 56 |
| Bacteroidetes | 3616 | 26206 | 14648 | 7270 | 2423 | 131835 |
| Chloroflexi | 0 | 6 | 1.333 | 2.236 | 0.7454 | 12 |
| Cyanobacteria | 0 | 52 | 18.67 | 19.87 | 6.625 | 168 |
| Elusimicrobia | 0 | 54 | 10 | 18.28 | 6.092 | 90 |
| Fibrobacteres | 0 | 2 | 0.4444 | 0.8819 | 0.294 | 4 |
| Firmicutes | 1400 | 101636 | 26578 | 41397 | 13799 | 239201 |
| Fusobacteria | 0 | 6 | 1.667 | 1.871 | 0.6236 | 15 |
| LD1 | 0 | 22 | 4.222 | 8.09 | 2.697 | 38 |
| Lentisphaerae | 0 | 2012 | 517.1 | 799.3 | 266.4 | 4654 |
| Planctomycetes | 0 | 42 | 12 | 14.63 | 4.876 | 108 |
| Proteobacteria | 26096 | 276648 | 89849 | 79417 | 26472 | 808644 |
| SR1 | 0 | 4 | 0.6667 | 1.414 | 0.4714 | 6 |
| Spirochaetes | 0 | 330 | 51.11 | 107.9 | 35.97 | 460 |
| TM7 | 4 | 840 | 160.7 | 281.9 | 93.96 | 1446 |
| Tenericutes | 2 | 340 | 63.89 | 111.8 | 37.28 | 575 |
| Verrucomicrobia | 8 | 1368 | 479.6 | 553.5 | 184.5 | 4316 |
| WPS-2 | 0 | 4 | 1.333 | 1.732 | 0.5774 | 12 |
| Thermi | 0 | 4 | 0.4444 | 1.333 | 0.4444 | 4 |