
Molecular analysis of the human faecal archaea in a southern Indian population

SANDYA B RANI, RAMADASS BALAMURUGAN and BALAKRISHNAN S RAMAKRISHNA*
The Wellcome Trust Research Laboratory, Christian Medical College, Vellore 632 004, India

*Corresponding author (Email, wurama@hotmail.com)

Archaea are an important constituent of the human gut microbiota, but there is no information on human gut archaea in an Indian population. In this study, faecal samples were obtained from different age groups (neonatal babies, pre-school children, school-going children, adolescents, adults and elderly) of a southern Indian population, and from a tribal population also resident in southern India). 16S rRNA gene sequences specific to Archaea were amplified from pooled faecal DNA in each group, sequenced, and aligned against the NCBI database. Of the 806 adequate sequences in the study, most aligned with 22 known sequences. There were 9 novel sequences in the present study. All sequences were deposited in the GenBank nucleotide sequence database with the following accession numbers: KF607113 - KF607918. *Methanobrevibacter* was the most prevalent genus among all the age groups accounting for 98% in neonates, 96% in post-weaning, and 100% each in preschool, school and adult population. In the elderly, *Methanobrevibacter* accounted for 96% and in tribal adults, 99% of the clones belonged to *Methanobrevibacter* genus. Other genera detected included *Caldisphaera*, *Halobaculum*, *Methanosphaera* and *Thermogymnomonas*. *Methanobrevibacter smithii* predominated in all age groups, accounting for 749 (92.9%) of the 806 sequences. Archaea can be found in the faeces of southern Indian residents immediately after birth. *Methanobrevibacter smithii* was the dominant faecal archeon in all age groups, with other genera being found at the extremes of age.

[Rani SB, Balamurugan R and Ramakrishna BS 2017 Molecular analysis of the human faecal archaea in a southern Indian population. *J. Biosci.* 42 113–119]

1. Introduction

The microbiota of the human gastrointestinal tract play major roles in human nutrition and metabolism (Ramakrishna 2013). While a large number of gut microbial species, including bacteria, archaea and fungi, have been cultured (Rajilić-Stojanović and de Vos 2014), there remain many gut microbes that are uncultivable. The enormous progress of the last two decades in characterizing the composition and function of the gut microbiota was primarily driven by the development of molecular tools for analysis and the development of molecular phylogenetics (Ramakrishna 2007; Lagier *et al.* 2012).

Archaea are single-celled microorganisms, distinct from bacteria, that do not have a cell nucleus or membrane-bound organelles within the cell, and are classified as a separate Domain or Kingdom. There are four phyla of archaea that

are widely distributed in the environment (Schleper *et al.* 2005; Auguet *et al.* 2010; Baker *et al.* 2010). Examination of 16S ribosomal RNA genes provides a powerful tool to study the gastrointestinal microbiota without culturing them (Yarza *et al.* 2014). Although they resemble bacteria morphologically, archaea possess genes and metabolic pathways (such as transcription and translation) that are found only in eukaryotes. One of the important biological properties of archaea is their ability to generate methane (Forterre *et al.* 2002; Cavicchioli 2011). Methanogenesis is important for preventing the accumulation of acids and reaction end-products in the human colon (McNeil 1984; Samuel *et al.* 2007).

The human faecal archaea have been characterized in Western populations, but no information is available on the human faecal archaea in Indians. Like the other gut microbiota, archaea are majorly influenced by dietary and

Keywords. Gut microbiota; methanogen; microbial diversity

environmental influences, and it is conceivable that there may be differences between populations with respect to the archaea present in stool. We undertook a study to identify the dominant archaea present in the faeces of southern Indian residents at different ages.

2. Methods

2.1 Samples

Faecal specimens, collected from healthy free living individuals in the community, were used for these studies. The studies were carried out in residents of semi-urban and rural areas in Vellore district of Tamil Nadu. Faecal samples from individuals of various ages ranging from children to the elderly were collected as described earlier (Balamurugan *et al.* 2008). Faecal samples from neonates were collected from the maternity ward of a hospital as described earlier (Kabeerdoss *et al.* 2013). Informed written consent was obtained for faecal sample collection and molecular analysis from adult participants, and from any one of the parents in the case of children. Faecal samples were received in the laboratory and stored at -80°C . All study participants were healthy and none had received any antibiotics in the 2 months prior to faecal collection. The study groups consisted of the following: neonates on the second day after birth; post-weaning children aged 6 months to 2 years; preschool children aged 3 years to 5 years; school-going children aged 6 years to 17 years; adults aged 18 years to 65 years; elderly individuals aged over 65 years; adult tribals 18 to 65 years residing in the Yelagiri hills.

2.2 Extraction and purification of DNA

DNA was extracted from 220 mg of faecal samples using the QIAamp® DNA Stool Mini Kit (Qiagen, Tilden, Germany), following the manufacturer's protocol, and eluted in 100 μL of buffer AE provided in the kit and stored at -20°C . DNA from 10 individuals in each group was pooled in equal amounts and formed the starting point for the following studies.

2.3 Archaeal 16S rRNA gene amplification

Segments of the 16S rRNA gene specific to Domain Archaea were amplified from the extracted DNA using the primers A333F (5'TCCAGGCCCTACGGG3') and A976R (5'YCCGGCGTTGAMTCCAATT3') (Baker and Cowan 2004). The segment of DNA amplified by the two primers includes a number of shorter sequences that cover up to 97% of Archaeal phyla and species (Wang and Qian 2009). DNA from the 10 individuals in each group (10 ng from each individual, 100 ng total) was pooled and used in the PCR. The mix contained 300 μM of each primer, 10 units of Ampliqon Taq DNA polymerase 2X Master Mix Red (0.2 units/ μL , 0.4 mM

dNTPs, 2 mM MgCl_2 final concentration). The PCR cycling was done for 40 cycles with initial denaturation at 95°C for 1 min followed by denaturation at 95°C for 30 s, 59°C annealing for 30 s, and 72°C for 2 min, followed by a 10 min 72°C extension. The amplicons thus obtained were checked on 1.0% agarose gel for amplification and the 607 bp band (figure 1) was excised and purified using QIAquick Gel Extraction kit (Cat.No. 28704; Qiagen, Tilden, Germany).

2.4 Cloning and plasmid isolation

PCR products were cloned into pCR®2.1-TOPO® vector (Invitrogen, Carlsbad, CA, USA), and One-Shot® MAX Efficiency® DH5 α ™-T1® chemically competent *E. coli* cells (Invitrogen, Carlsbad, CA, USA) were transformed using the TOPO-TA Cloning® kit (Invitrogen, Carlsbad, CA, USA), according to the manufacturer's instructions. Single white colonies were picked up and inoculated in 3 mL LB broth with ampicillin (75 $\mu\text{g}/\text{mL}$) and incubated overnight at 37°C at

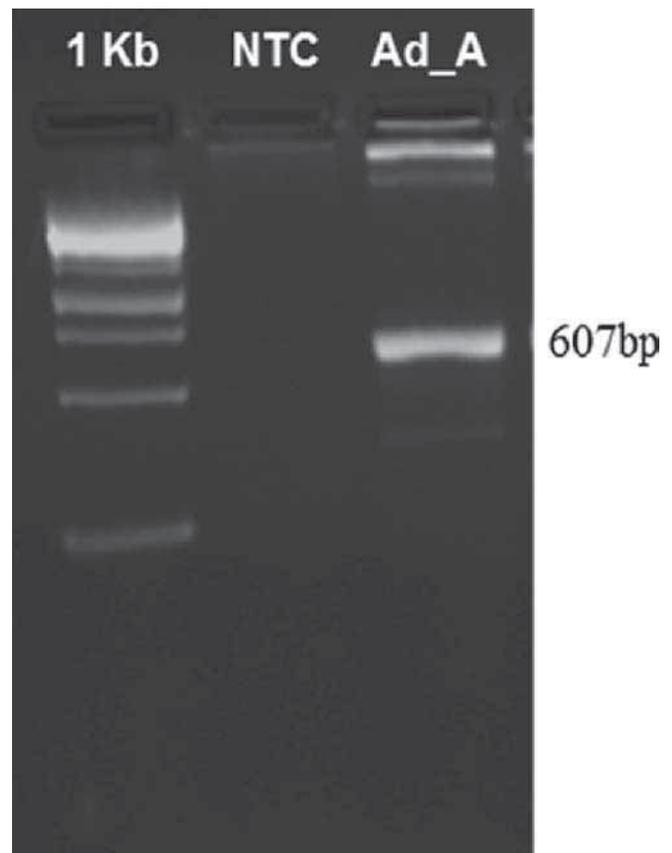


Figure 1. Representative gel picture showing PCR product after amplification using the primers in this study. Lane on the left shows the 1 kb DNA ladder, the middle lane has the non-template control (NTC), while the right lane shows the 607 bp product from the adult group (Sample_Ad).

260 rpm in a shaking incubator. 850 μ L of culture was added to 150 μ L of glycerol and stored at -80°C . 800 μ L of glycerol stock was taken and inoculated in 8 mL of LB broth with ampicillin and incubated overnight at 37°C , at 260 rpm. Plasmids were isolated using a plasmid isolation kit (Cat. No. PLN 350, Sigma-Aldrich, St Louis, MO, USA) and eluted in 75 μ L of sterile water. The plasmids were quantified by fluorimetry using SYBR green and stored at -20°C .

2.5 PCR crosscheck of clones

The inserts in the plasmids were cross checked by PCR amplification using the same set of primers (A333F and A976R). The plasmids which showed 607 bp product were considered as recombinant clones.

2.6 Sequencing of clones

120 plasmids from each group were sequenced using Big Dye terminator cycle sequencing kit v.3.1 (Applied Biosystems, Foster City, CA, USA) and M13F (5'GTAAAACGACGGCCAGT3') sequencing primers on an Applied Biosystems model 3730XL automated DNA sequencing system at Macrogen, Seoul, Korea.

2.7 Sequence analysis

120 clones from each age group were sequenced unidirectionally using M13F primer and examined to provide a comprehensive picture of the faecal archaeal flora. The raw sequence data from chromatogram was converted to FASTA format. The vector sequences were identified using Vec screen tool (NCBI) and vector sequences were removed manually. Vector trimmed sequences were checked for Chimeras using Database Enabled Code for Ideal Probe Hybridization Employing R (DECIPHER) software (Wright *et al.* 2012). No chimeras were detected in any of the sequences. Sequences with poor quality reads and short read lengths were excluded. Clean sequences were used for taxonomic identification and phylogenetic analysis. All the sequences were thereafter identified through the GenBank database (www.ncbi.nlm.nih.gov) using the BLAST (basic local alignment search tool) algorithm (Altschul *et al.* 1997) and the 'Classifier' tool of the Ribosomal Database Project (RDP) II (Cole *et al.* 2005). 806 sequences were submitted to GenBank through Bankit.

2.8 Phylogenetic analysis

The 806 sequences from the study and 22 sequences from NCBI to which our sequences corresponded were

downloaded in FASTA format from GenBank using MEGA 5.2.2. Sequences from the present study which matched available sequences from NCBI with $\geq 99\%$ similarity were considered to be identical to the NCBI sequences and were labelled with NCBI accession IDs. Sequences with less than 99% similarity were labelled with individual clone names. Thirty-one unique sequences were uploaded to MEGA 5.2 software. Clustal W program with default parameters computed the pairwise alignment and multiple sequence alignment. Pairwise distances were calculated for the aligned sequences. Overall mean distance was 0.4 and sequences were used to produce an unrooted phylogenetic tree according to the neighbour-joining method with 1000 bootstrap replicates and using the Jukes-Cantor model.

2.9 Nucleotide sequence accession numbers

The nucleotide sequences reported in this study appear in the GenBank nucleotide sequence database with the following accession numbers: Ne_A - KF607113 - KF607228; PW_A - KF607229 - KF607339; PS_A - KF607567 - KF607681; SG_A - KF607340 - KF607449; Ad_A - KF607450 - KF607566; El_A - KF607682 - KF607798; and TA_A - KF607799 - KF607918.

2.10 Ethical considerations

The consent forms and study protocol were approved by the Institutional Review Board (IRB), Christian Medical College, Vellore.

3. Results

Cleaned sequences from neonate (116 clones), post-weaning (111 clones), preschool (115 clones), school going (110 clones), rural adult (117 clones), elderly population (117 clones) and tribal adult (120 clones) were used for taxonomic and phylogenetic analysis. The sequence length of the evaluated clones is shown in figure 2.

3.1 Taxonomic analysis

Of the 806 sequences in the present study, many corresponded to 22 sequences already recorded in the NCBI database. Nine novel sequences in the present study did not find appropriate matches in the NCBI database and have been uploaded in GenBank. Archaea belonging to 2 phyla and 5 genera were detected in the different groups. Euryarchaeota and Crenarchaeota were the two phyla detected. Euryarchaeota formed the most abundant phylum. *Methanobrevibacter* was the most prevalent genus among all the age groups accounting for 98% in neonates, 96% in

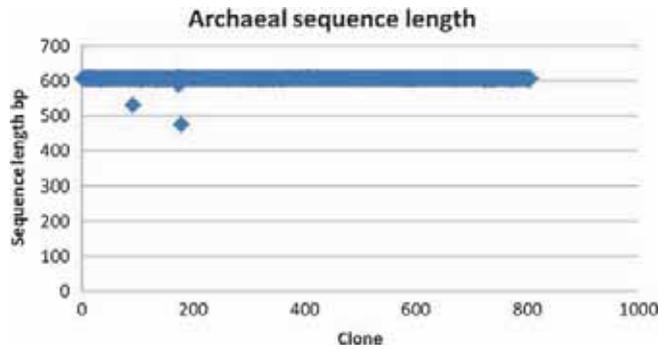


Figure 2. Length distribution of the amplicons that were cloned and sequenced.

post-weaning, and 100% each in preschool, school-going and adult population. In the elderly, *Methanobrevibacter* accounted for 96% and in tribal adults, 99% of the clones belonged to *Methanobrevibacter* genus. Other genera detected in very minor proportions belonged to *Caldisphaera* (Phylum Crenarchaeota), *Halobaculum* (Phylum Euryarchaeota), *Methanosphaera* (Phylum Euryarchaeota), and *Thermogymnomonas* (Phylum Euryarchaeota) (table 1).

Figure 3 depicts the distribution of archaeal sequences in the faeces of the different age groups that were studied. As can be seen from this, the overwhelming majority of faecal archaeal sequences in all the groups belonged to *Methanobrevibacter smithii*.

3.2 Distribution of archaeal taxa among the different study groups

We attempted to compare the distribution of archaeal taxa among the various study groups. The post-weaning children, school children, and adolescent groups were clubbed, as were the tribal and rural adult groups, and the neonate, children, and elderly were compared. Venn diagrams were generated to show the overlap and differences between these four groups (figure 4).

3.3 Phylogeny

The 22 unique NCBI-matched sequences and the 9 novel sequences identified in this study were used to construct an unrooted phylogenetic tree (figure 5). The novel archaeal sequences identified in the present study were clustered separately from the archaea already recorded in the database.

4. Discussion

We describe, for the first time, the faecal archaea of a southern Indian population, determined using a cultivation-independent molecular approach. The dominant presence of *Methanobrevibacter* in the faeces in all age groups as well as the presence of other archaea, including several novel ones, is described.

Many of the faecal archaea are methane producers or methanogens. The methanogenic archaeon *Methanobrevibacter smithii* was first identified in 1982 from a human faecal culture (Miller *et al.* 1982). Early studies used culture to identify methane-producing microbes and found that they were predominantly located in the right colon (Pochart *et al.* 1993). *Methanobrevibacter smithii* has been reported to be the most prevalent archaeal species in human faeces followed by *Methanosphaera stadtmanae*, which was, however, much less common (Dridi *et al.* 2009; Mihajlovski *et al.* 2010). Non-methanogenic archaea including members of Thermoplasma, the Crenarchaeota and halophilic archaea have also been reported in the gastrointestinal tract. The literature provides conflicting information regarding the occurrence of the archaea in children. Methanogenic archaea could be detected in the faeces of children in New Zealand (Stewart *et al.* 2006), while they could not be detected in Italian children before the age of 27 months (Rutili *et al.* 1996). An increase in the diversity of methanogenic archaeal species in faeces was found in elderly individuals compared to adults (Dridi *et al.* 2011).

The present study found that *M. smithii* was the most prevalent archaeal species in faeces. It ranged from 88.7% in neonates and the elderly, to 97.2% in school-going children. *Methanosphaera* was the next most common archaeal genus in this study. We also

Table 1. Archaeal genera identified in each of the study groups

Group	<i>Methanobrevibacter smithii</i>	<i>Methanobrevibacter other</i>	Other genera
Neonatal	103	11	Methanosphaera 2
Post-weaned children	104	3	Methanosphaera 3 Halobaculum 1
Pre-school children	105	10	0
School going children	107	3	0
Adult	112	5	0
Tribal adult	114	5	Caldisphaera 1
Elderly	104	9	Thermogymnomonas 4

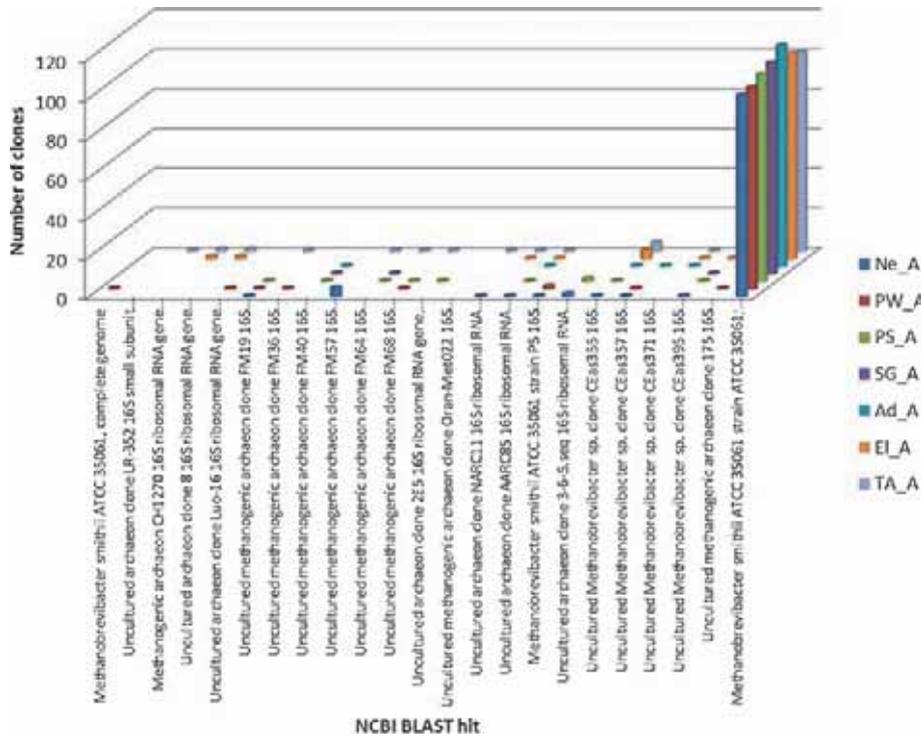


Figure 3. Frequency of different archaeal species among the fecal microbiota of the different groups of healthy individuals in this study. Ne_A, neonatal archaea; PW_A, pre-weaning children; PS_A, pre-school children; SG_A, school going children; Ad_A, adults; El_A, elderly; TA_A, tribal adults.

found *Thermogymnomonas*, *Halobaculum* and *Caldisphaera* although they were much less abundant than *Methanobrevibacter*. In addition, we identified several novel archaea sequences which have not been previously described and these sequences have been deposited in GenBank. Archaeal genera other than *Methanobrevibacter* were particularly noted at the extremes of age. Our studies suggest that increased diversity occurs at both ends of the age spectrum. Others have also noted an increased diversity of methanogenic archaea in elderly individuals as compared to adults (Dridi *et al.* 2011). This is very likely related to differences in dietary practices in these age groups.

Few studies have examined the presence of archaea in the gut microbiota of children. A study in Italy did not find methanogenic archaea in children below the age of 27 months (Rutili *et al.* 1996). A study in a Western population reported that archaea transiently occurred in the faeces of neonates in the first weeks of life and thereafter reappeared only at a much later age (Palmer *et al.* 2007). In the present study, faecal archaea could be identified in all age groups including newborn infants and children of all ages. It is possible that there may be differences in the age of appearance of archaea in the stool in developing and developed countries. The very early archaeal colonization of the

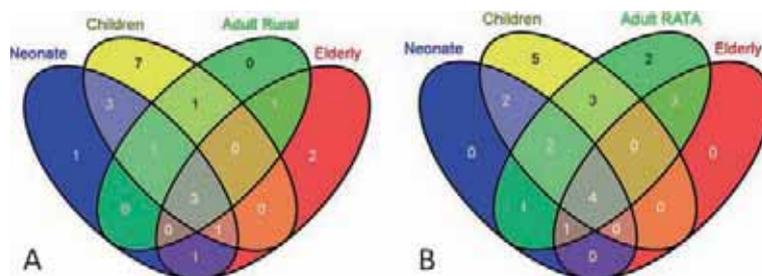


Figure 4. Venn diagrams showing distribution of the archaeal taxa in different age groups. Panel (A) includes tribal population. Four-way Venn diagrams were plotted in (A) comparing neonate, children (PW, PS and SG), rural adults and elderly using VENNY (<http://bioinfo.gp.cnb.csic.es/tools/venny/index.html>). In Panel (B) the addition of tribal adults to the rural adult pool showed that some of the neonate and children archaea were present in the tribal adults.

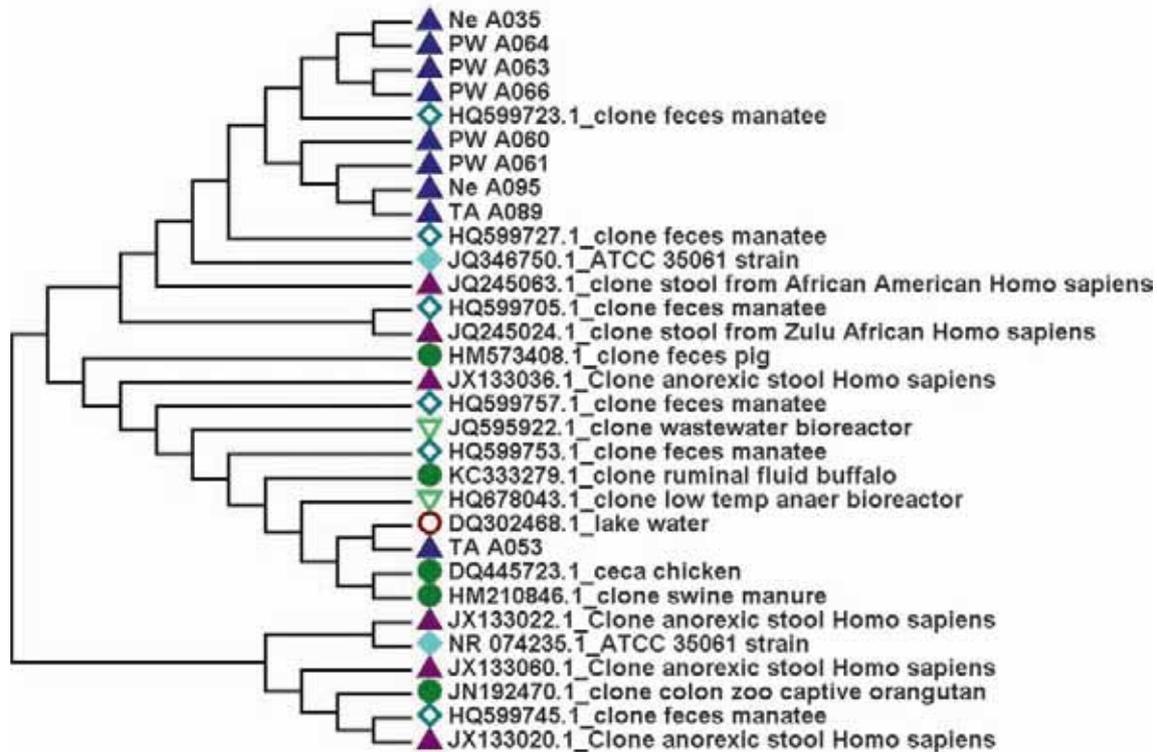


Figure 5. Phylogenetic tree constructed from 31 unique sequences identified in the present study. Twenty-two sequences that corresponded with >99% similarity to those already available in the NCBI database were labelled with the NCBI accession number, while nine sequences that were described for the first time in this study are depicted by green diamonds and marked with study identifier numbers (Ne, neonate; PW, pre-weaning; TA, tribal adult).

gut in the developing countries is consistent with the very early microbial colonization of the gut that we have earlier noted in newborn infants in southern India (Kabeerdoss *et al.* 2013).

Several of the human faecal archaea reported in the present study matched to strains that were originally described as environmental strains or animal strains. We have earlier noted gut colonization of humans by animal microbiota in individuals living in close proximity to animals (Balamurugan *et al.* 2009).

The significance of the archaeal population of the gut in human health and disease has not been studied to the same extent to which the bacterial contribution has been studied. *Methanobrevibacter* produces methane from by-products of bacterial fermentation in the gut. It contributes to human nutrition and metabolism (Vanderhaeghen *et al.* 2015). Recent studies have begun to identify alterations of the archaea in human disease. It has been reported that there is a reduction of methanogenic archaea in patients with irritable bowel syndrome (Pozuelo *et al.* 2015) while *Methanosphaera stadtmanae* abundance has been reported to increase in inflammatory bowel disease (Blais Lecours *et al.* 2014).

The strength of the study is that it was based on full-length cloning and sequencing of the archaeal 16S rRNA

gene. There are three limitations of the study: First, we used unidirectional sequencing since our sequencing strategy provided usable reads up to 700 bp. It is possible that bidirectional sequencing would have provided a marginal increase in usable sequences. Second, the depth of sequencing was much less than that which can be achieved by next generation sequencing; however, the latter results in short length sequences which then require to be assembled using appropriate strategies while the current strategy provided the exact sequence without the need for assembly. Third, the samples were pooled and not individually sequenced; again, this was not considered important in the context of discovery. Although these limitations are acknowledged, this is the first report of the faecal archaea in an Indian population and is likely to be of use when studies are designed to evaluate the contribution of the gut archaea in human health in Indians.

Acknowledgements

This study was funded by grant no. BT/PR-10769/FNS/20/384/2008 from the Department of Biotechnology, New Delhi, India.

References

- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25** 3389–3402
- Auguet JC, Barberan A and Casamayor EO 2010 Global ecological patterns in uncultured Archaea. *ISME J.* **4** 182–190
- Baker GC and Cowan DA 2004 16S rDNA primers and the unbiased assessment of thermophile diversity. *Biochem. Soc. Trans.* **32** 281–221
- Baker BJ, Comolli LR, Dick GJ, Hauser LJ, Hyatt D, Dill BD, *et al.* 2010 Enigmatic, ultrasmall, uncultivated Archaea. *Proc. Natl. Acad. Sci. USA* **107** 8806–8811
- Balamurugan R, Janardhan HP, George S, Chittaranjan SP and Ramakrishna BS 2008 Bacterial succession in the colon during childhood and adolescence: molecular studies in a southern Indian village. *Am. J. Clin. Nutr.* **88** 1643–1647
- Balamurugan R, Aarthi CM, Chittaranjan SP and Ramakrishna BS 2009 Molecular detection of the ruminal bacterium, *Butyrivibrio fibrisolvens*, in feces from rural residents of southern India. *Microb. Ecol. Health Dis.* **21** 38–43
- Blais Lecours P, Marsolais D, Cormier Y, Berberi M, Haché C, Bourdages R and Duchaine C 2014 Increased prevalence of *Methanospaera stadmanae* in inflammatory bowel diseases. *PLoS One* **9** e87734
- Cavicchioli R 2011 Archaea—timeline of the third domain. *Nat. Rev. Microbiol.* **9** 51–61
- Cole JR, Chai B, Farris RJ, Wang Q, Kulam SA, McGarrell DM, Garrity GM and Tiedje JM 2005 The Ribosomal Database Project (RDP-II): sequences and tools for high-throughput rRNA analysis. *Nucleic Acids Res.* **33** D294–D296
- Dridi B, Henry M, El Khechine A, Raoult D and Drancourt M 2009 High prevalence of *Methanobrevibacter smithii* and *Methanospaera stadmanae* detected in the human gut using an improved DNA detection protocol. *PLoS One* **4** e7063
- Dridi B, Raoult D and Drancourt M 2011 Archaea as emerging organisms in complex human microbiomes. *Anaerobe.* **17** 56–63
- Forterre P, Brochier C and Philippe H 2002 Evolution of the Archaea. *Theor. Popul. Biol.* **61** 409–422
- Kabeerdoss J, Ferdous S, Balamurugan R, Mechenro J, Vidya R, Santhanam S, Jana AK and Ramakrishna BS 2013 Development of the gut microbiota in southern Indian infants from birth to six months: a molecular analysis. *J Nutr Sci.* **2**, e18
- Lagier JC, Million M, Hugon P, Armougom F and Raoult D 2012 Human gut microbiota: repertoire and variations. *Front. Cell Infect. Microbiol.* **2** 136
- McNeil NI 1984 The contribution of the large intestine to energy supplies in man. *Am J Clin Nutr.* **39** 338–342
- Mihajlovski A, Dore J, Levenez F, Alric M and Brugere JF 2010 Molecular evaluation of the human gut methanogenic archaeal microbiota reveals an age-associated increase of the diversity. *Environ. Microbiol. Rep.* **2** 272–280
- Miller TL, Wolin MJ, Conway de Macario E and Macario AJ 1982 Isolation of *Methanobrevibacter smithii* from human feces. *Appl. Environ. Microbiol.* **43** 227–232
- Palmer C, Bik EM, DiGiulio DB, Relman DA and Brown PO 2007 Development of the human infant intestinal microbiota. *PLoS Biol.* **5** e177
- Pochart P, Lemann F, Flourie B, Pellier P, Goderel I and Rambaud JC 1993 Pyxigraphic sampling to enumerate methanogens and anaerobes in the right colon of healthy humans. *Gastroenterology* **105** 1281–1285
- Pozuelo M, Panda S, Santiago A, Mendez S, Accarino A, Santos J, Guarner F, Azpiroz F, *et al.* 2015 Reduction of butyrate- and methane-producing microorganisms in patients with Irritable Bowel Syndrome. *Sci. Rep.* **5** 12693. doi:10.1038/srep12693
- Rajilić-Stojanović M and de Vos WM 2014 The first 1000 cultured species of the human gastrointestinal microbiota. *FEMS Microbiol. Rev.* doi:10.1111/1574-6976.12075
- Ramakrishna BS 2007 The normal bacterial flora of the human intestine and its regulation. *J. Clin. Gastroenterol.* **41** S2–S6
- Ramakrishna BS 2013 Role of the gut microbiota in human nutrition and metabolism. *J. Gastroenterol. Hepatol.* **28** 9–17
- Rutili A, Canzi E, Brusa T and Ferrari A 1996 Intestinal methanogenic bacteria in children of different ages. *New Microbiol.* **19** 227–243
- Samuel BS, Hansen EE, Manchester JK, Coutinho PM, Henrissat B, Fulton R, *et al.* 2007 Genomic and metabolic adaptations of *Methanobrevibacter smithii* to the human gut. *Proc. Natl. Acad. Sci. USA* **104** 10643–10648
- Schleper C, Jurgens G and Jonuscheit M 2005 Genomic studies of uncultivated archaea. *Nat. Rev. Microbiol.* **3** 479–488
- Stewart JA, Chadwick VS and Murray A 2006 Carriage, quantification, and predominance of methanogens and sulfate-reducing bacteria in faecal samples. *Lett. Appl. Microbiol.* **43** 58–63
- Vanderhaeghen S, Lacroix C and Schwab C 2015 Methanogen communities in stools of humans of different age and health status and co-occurrence with bacteria. *FEMS Microbiol. Lett.* **362**, fmv092. doi:10.1093/femsle/fmv092
- Wang Y and Qian PY 2009 Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. *PLoS One* **4** e7401
- Wright ES, Yilmaz LS and Noguera DR 2012 DECIPHER, a search-based approach to chimera identification for 16S rRNA sequences. *Appl. Environ. Microbiol.* **78** 717–725
- Yarza P, Yilmaz P, Priesse E, Glöckner FO, Ludwig W, Schleifer KH, Whitman WB, Euzéby J, *et al.* 2014 Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat. Rev. Microbiol.* **12** 635–645

MS received 20 August 2016; accepted 17 January 2017