**Strain level establishment of a human core microbiota.** By Melissa Schanche[1], Ekaterina Avershina[1], Christian Dotterud[2], Torbjørn Øien[2], Ola Storrø[2], Roar Johnsen[2], and Knut Rudi[1]. *[1]Department of Chemistry, Biotechnology and Food Science, University of Life Sciences, Ås, Norway, [2]Department of Public Health and General Practice, Norwegian University of Science and Technology, Trondheim, Norway*

## Abstract

We therefore developed a novel Illumina deep sequencing approach targeting the 6S–23S rRNA Internal Transcribed Spacer (ITS) region. Using this approach we analyzed the strain level composition of the microbiota in a longitudinal cohort of 20 mother child pairs, where 10 of the mothers received probiotics during late pregnancy and till 90 days after delivery. We found that the probioics had a minor strain level impact on the mothers fecal, and breast milk microbiota, and the infants fecal microbiota. We found, however, a significant evidence for a core microbiota in both mother's breast- milk and stool, with *Streptococcus gordonii* as the most widespread colonizer. *S. gordonii* is also known as the most widespread dental colonizer. Although a core microbiota was not yet established at 10 - and 90 days post-delivery, did our data indicate that core strains were transmitted from mother to child with age through breast-milk. In conclusion, our data support the presence of a human core microbiota with cosmopolitan colonizers providing potential benefits to the host, as in the concept of the recently proposed interface hypothesis.

## Introduction

Taxonomic assignments of complex microbial assemblages have until now mainly been performed using 16S rRNA gene analyses. A major limitation with these analyses, however, is that taxonomic assignments are mainly restricted to the genus level (1). This limits the possibilities to address questions about prevalence of and transmission on strain level.

The establishment of the human gut microbiota can be considered the result of a succession process (2). Numerous studies have been conducted to intervene with this establishment, with the results from these interventions being partly conflicting (3). Strain level transmission patterns of both commensal, and probiotic bacteria have not yet been properly addressed (2, 4). Mother's milk and feces have been proposed as important vectors for transmission of bacteria from mother to child (5). Strain level information, however, is pivotal for describing the origin and prevalence of these bacteria (6).

Here we describe the strain level composition of the microbota in mothers and children from a double blinded randomized longitudinal trial of probiotic administration (7). Strain level information was obtained using a novel Illumina 16S–23S rRNA Internal Transcribed Spacer (ITS) region deep sequencing approach. This region has previously been extensively used for strain level characterization, utilizing a range of detection approaches. To our knowledge, however, the ITS region has not yet been explored for deep sequencing applications.

We found a surprisingly stable core of strains associated with both the mother's milk and stool samples – being resistant to the probiotic administration. The infant stool, however, did not show an established core of strains. We therefore propose that the core microbiota is transmitted and

established with age. These results are discussed in light of general hypotheses about determinants of gut microbiota composition (8-10).

## Materials and methods

**Clinical samples** The clinical samples represent a subset of the Pro-PACT study, which is a randomized double blinded study of the effect of giving mother's probiotics on the development of allergy for the child (7). From the study cohort we randomly selected 10 mother/child pairs which were given probiotics, and 10 pairs that were given placebo. Five samples were selected for each mother/child pair. For the mothers, breast milk samples from 10 and 90 days after delivery, in as well as stool samples from 90 days, were analyzed. For the children, stool samples were analyzed 10 and 90 days after delivery.

**PCR amplification and Illumina sequencing** For the 16S-23S rRNA ITS PCR we used the 16S IX-23S I primer pair with denaturation, annealing and extension for one minute each. The 16S rRNA gene was amplified using the general PRK primers, with the original protocol. The sequencing was done using the MiSeq sequencing platform (Illumina, San Diego, California, USA ) with the Reagents Kit v3 (300 bp paired end sequencing).

## Results

**Overall microbiota composition** We detected in total 1356 OTU's (corresponding to approximately 340 strains), represented by two or more sequences. The three most dominant OTU's, OTU88 (*Streptococcus gordonii*), OTU349 (*Streptococcus agalactiae*) and OTU375 (*Streptococcus pneumoniae*), represented 9.5 %, 6.6 % and 6.6 % of the sequences, respectively. The highest level of colonization by these OTU's was in mother's milk, where there seemed to be about equal amounts of the three species (Fig. 1A). However, in both mother and infant stool samples the level of *S. gordonii* was significantly higher ($p < 0.01$, Mann-Whitney U test) than of both *S. agalactiae* and *S. pneumoniae* (Fig. 1B and C).

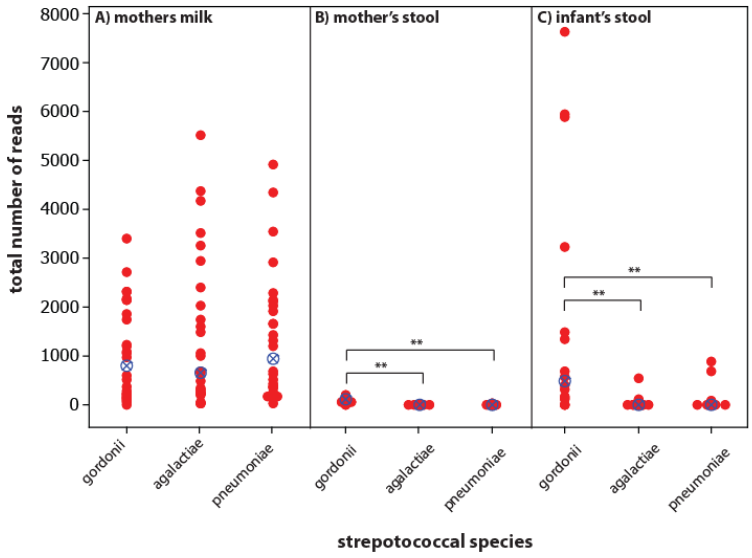


**Figure 1.Distribution of the three most abundant species in (A) mother's milk, (B) mothers stool and (C) infants stool.** The number of reads per sample (out of 10000) are indicated with red dots, while the means are indicated with the crossed dots. Significant differences are indicated with ** ( $p < 0.01$ Mann-Whitney U test)

For the overall composition of the microbiota we found that there was a slight overlap in the microbiota between mother's milk and the child's feces, while the mother's fecal samples formed a sub-cluster within the child's fecal samples. Strains belonging to the

streptococci were associated with mother's milk, while the child's feces and mothers feces were associated with a higher number of strains.

**Strain level prevalence** The prevalence of OTU's (detected more than once in a given sample) for the infant stool samples followed a negative geometric distribution (Fig. 2 A and B). However, for the mother's sample we found a deviation from the negative geometric distribution for the highly prevalent OTU's (present in more than 60 % of the samples). The most prevalent OTU's were clearly overrepresented ($p < 0.0005$ t-test deviation from the respective regression lines) compared to the negative geometric distribution. Furthermore, for the prevalent OTU's there were positive associations between the prevalence and number of OTU's (Fig. 2C to E).

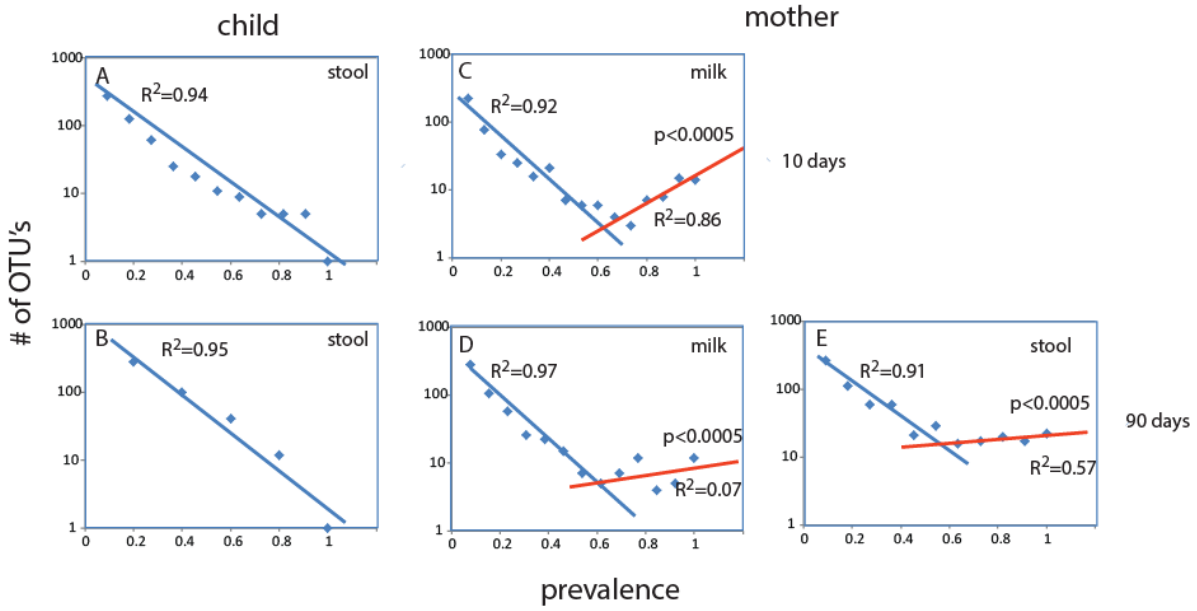


**Figure 2. Number of OTUs at different prevalence levels.** The regression lines represent fitting to a geometric distribution. Blue lines represent fitting of OTU's with prevalence's below 60%, while the red lines represent fitting to OTU's above. Since no OTU's were identified in all samples for the children we used one to have a data point. The p value indicates the deviation of the number of the most prevalent OTU's from the distribution below 60%.

For the mother's milk samples the most prevalent OTU's belonged to streptococci, while most of the highly prevalent OTU's from the mothers stool samples belonged to the clostridia.

## Discussion

Our results support the presence of a strain level core microbiota, both in the mothers stool and milk samples. We have previously identified strains belonging to the *Lachnospiraceae* as the core microbiota in the adult gut (9). The stable strains identified in the mother's stool mainly fall within clostridia, but unfortunately due to lack of reference strains the further classification remains uncertain for the 16S-23S rRNA ITS. Due to higher number of reference strains for the mother's milk, we obtained a better assignment for the stable strains in it. Here, an OTU belonging *S. gordonii* was the most abundant and prevalent. This species is a highly cosmopolitan colonizer of the human body, also being an important dental colonizer (11). This bacterium can produce hydrogen peroxide – potentially preventing the colonization by pathogens

(12). Furthermore, despite the high numbers and prevalence, relatively few incidences of infections have been reported for this bacterium (11). The low pathogenicity is also confirmed by that only few virulence genes (about 70) compared to more than 250 for the closely related *S. pneumonia* (identified from the Victors database), were detected in its genome. We therefore favor that *S. gordonii* bacterium has a positive and not an opportunistic host association. Although *S. agalactiae* and *S. pneumoniae* showed about equal colonization of mother's milk, they had a much lower colonization of both infant and mothers stool than *S. gordonii*, suggesting that *S. gordonii* is allowed in more body sites.

The currently most widely accepted hypothesis for host-bacterial interactions state the absence of a phylogenetic core, and no direct host selection (13). We have, however, recently proposed the interface hypothesis as an alternative explanation, stating that service providing bacteria have a close host association (8).

The results presented here are in favor of the interface hypothesis for the following reasons: a) Based on the prevalence data we found strong evidence for a phylogenetic core microbiota. b) The most cosmopolitan colonizer *S. gordonii* has low virulence species is allowed in multiple body sites, suggesting a close host association. c) The cosmopolitan colonizer *S. gordonii* provides a potential host service through hydrogen peroxide production and pathogen protection.

## References


1. Baker GC, Smith JJ, Cowan DA. Review and re-analysis of domain-specific 16S primers. J Microbiol Methods. 2003;55(3):541-55.
2. Avershina E, Storro O, Oien T, Johnsen R, Pope P, Rudi K. Major faecal microbiota shifts in composition and diversity with age in a geographically restricted cohort of mothers and their children. FEMS microbiology ecology. 2013.
3. Sanz Y. Gut microbiota and probiotics in maternal and infant health. The American journal of clinical nutrition. 2011;94(6 Suppl):2000S-5S.
4. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, et al. Human gut microbiome viewed across age and geography. Nature. 2012;486(7402):222-7.
5. Jeurink PV, van Bergenhenegouwen J, Jimenez E, Knippels LM, Fernandez L, Garssen J, et al. Human milk: a source of more life than we imagine. Benef Microbes. 2013;4(1):17-30.
6. Sanders ME, Guarner F, Guerrant R, Holt PR, Quigley EM, Sartor RB, et al. An update on the use and investigation of probiotics in health and disease. Gut. 2013;62(5):787-96.
7. Dotterud CK, Storro O, Johnsen R, Oien T. Probiotics in pregnant women to prevent allergic disease: a randomized, double-blind trial. Br J Dermatol. 2010;163(3):616-23.
8. Avershina E, Rudi K. Is it who you are or what you do that is important in the human gut? Benef Microbes. 2013;4(3):219-22.
9. Sekelja M, Berget I, Naes T, Rudi K. Unveiling an abundant core microbiota in the human adult colon by a phylogroup-independent searching approach. Isme Journal. 2011;5(3):519-31.
10. Magurran AE, Henderson PA. Explaining the excess of rare species in natural species abundance distributions. Nature. 2003;422(6933):714-6.
11. Lee SF. Oral colonization and immune responses to Streptococcus gordonii: Potential use as a vector to induce antibodies against respiratory pathogens. Curr Opin Infect Dis. 2003;16(3):231-5.
12. Zheng L, Itzek A, Chen Z, Kreth J. Environmental influences on competitive hydrogen peroxide production in Streptococcus gordonii. Applied and environmental microbiology. 2011;77(13):4318-28.
13. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. Diversity, stability and resilience of the human gut microbiota. Nature. 2012;489(7415):220-30.