The gut as reservoir of antibiotic resistance
microbial diversity of tetracycline resistance in mother and infant

de Vries, Lisbeth Elvira; Vallès, Yvonne; Agersø, Yvonne; Vaishampayan, Parag A.; Garcia-Montaner, Andrea; Kuehl, Jennifer V.; Christensen, Henrik; Barlow, Miriam; Francino, M. Pilar

Published in:
P L o S One

DOI:
10.1371/journal.pone.0021644

Publication date:
2011

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
The Gut as Reservoir of Antibiotic Resistance: Microbial Diversity of Tetracycline Resistance in Mother and Infant

Lisbeth E. de Vries¹, Yvonne Vallès³, Yvonne Agero², Parag A. Vaishampayan⁴, Andrea García-Montaner³, Jennifer V. Kuehl⁴, Henrik Christensen¹, Miriam Barlow⁵, M. Pilar Francino³,⁴, ⁵

¹ Department of Veterinary Disease Biology, University of Copenhagen, Frederiksberg, Denmark, ² National Food Institute, Technical University of Denmark, Lyngby, Denmark, ³ Unitat Mixta d’Investigació en Genòmica i Salut, Centre Superior d’Investigació en Salut Pública/Universitat de València-Institut Cavanilles, València, Spain, ⁴ Evolutionary Genomics Program, Department of Energy Joint Genome Institute, Walnut Creek, California, United States of America, ⁵ School of Natural Sciences, University of California Merced, Merced, California, United States of America

Abstract

The microbiota in the human gastrointestinal tract (GIT) is highly exposed to antibiotics, and may be an important reservoir of resistant strains and transferable resistance genes. Maternal GIT strains can be transmitted to the offspring, and resistances could be acquired from birth. This is a case study using a metagenomic approach to determine the diversity of microorganisms conferring tetracycline resistance (TcR) in the guts of a healthy mother-infant pair one month after childbirth, and to investigate the potential for horizontal transfer and maternal transmission of TcR genes. Fecal fosmid libraries were functionally screened for TcR, and further PCR-screened for specific TcR genes. TcR fosmid inserts were sequenced at both ends to establish bacterial diversity. Mother and infant libraries contained TcR, although encoded by different genes and organisms. TcR organisms in the mother consisted mainly of Firmicutes and Bacteroidetes, and the main gene detected was tet(O), although tet(W) and tet(X) were also found. Identical TcR gene sequences were present in different bacterial families and even phyla, which may indicate horizontal transfer within the maternal GIT. In the infant library, TcR was present exclusively in streptococci carrying tet(M), tet(L) and erm(T) within a novel composite transposon, Tn6079. This transposon belongs to a family of broad host range conjugative elements, implying a potential for the joint spread of tetracycline and erythromycin resistance within the infant’s gut. In addition, although not found in the infant metagenomic library, tet(O) and tet(W) could be detected in the uncloned DNA purified from the infant fecal sample. This is the first study to reveal the diversity of TcR bacteria in the human gut, to detect a likely transmission of antibiotic resistance from mother to infant GITs and to indicate the possible occurrence of gene transfers among distantly related bacteria cohabiting the GIT of the same individual.

Introduction

The human gastrointestinal tract (GIT) is host to a very dense microbiota, harboring $10^{13}$–$10^{14}$ bacterial cells in adults and a broad diversity of bacterial species, of which a large proportion are not yet cultured. This microbiota is often exposed to a variety of antibiotics, both directly and indirectly, due to their routine use in clinical settings and in farm animals. Therefore, its many other fundamental roles in health notwithstanding [1–5], the GIT microbiota may serve as an important reservoir of antibiotic resistant strains that could act as opportunistic pathogens or as donors of resistance genes to other bacteria [6]. In infants, infections due to antibiotic resistant strains are on the rise and represent a major cause of mortality and morbidity worldwide. Although the infant’s gut is thought to be mostly germ-free at birth, it rapidly enters an extensive and complex process of colonization by a variety of microbes [7,8], and recent studies have firmly established that strains from the mother’s GIT can be transmitted to the infant and persist during the first weeks of life [9]. Consequently, antibiotic resistances could be vertically transmitted from the maternal GIT and bear on infant health from a very early age.

Tetracyclines are one of the most widely used groups of antibiotics worldwide and tetracycline resistance (TcR) is extremely common among bacteria [10]. Presently, 43 distinct TcR genes are known and they are usually associated with large mobile genetic elements (MGE) (http://faculty.washington.edu/marilyn/). The most common forms are the active efflux of tetracycline from the cell and the synthesis of ribosomal protection proteins that prevent the binding of tetracycline to the ribosomes [10,11]. Although its medical applications have decreased in the last decade and it is no longer used for treatment of pregnant women or children under...
the age of 8 years [12], tetracycline is still widely used for therapeutic treatment in animal production and in some countries it is also used as growth promoter in animal feed [11,13]. Therefore intestinal bacteria are still extensively exposed to this antibiotic.

A recent microarray-based study has found tet(M) and tet(W) to be the most prevalent Tc\(^r\) genes for the oral and fecal metagenomes of healthy adults, respectively [14]. Furthermore, Tc\(^r\) genes like tet(M), tet(O) and tet(W) have also been detected in fecal samples from healthy and exclusively breast-fed infants, suggesting that Tc\(^r\) genes are common in the environment [15]. However, these studies have not revealed the types of bacteria that harbor these resistances in the GIT and have not addressed the potential origin of the Tc\(^r\) genes and strains present in healthy infants. Here, we have used a culture-independent approach to characterize the diversity of microorganisms conferring Tc\(^r\) in the gut of one healthy infant-mother pair. Two fecal metagenomic libraries, one from the mother and one from her exclusively breast-fed infant one month after birth [9], were screened for ribosomal protection genes none of the assayed Tc\(^r\) genes were detected, and all clones were negative for \(\text{erm}(T)\), \(\text{tet}(S)\), \(\text{tet}(G)\), \(\text{tet}(L)\), \(\text{tet}(A)\), \(\text{tet}(P)\), \(\text{tet}(Q)\) and \(\text{erm}(T)\).

**Results**

**Screening metagenome libraries from infant and mother for clones conferring Tc\(^r\)**

The metagenome libraries from the infant and the mother contained 44 and 272 fosmid Tc\(^r\) clones respectively. In a first instance, we screened all obtained Tc\(^r\) clones for the common ribosomal protection genes \(\text{tet}(M)\), \(\text{tet}(O)\), \(\text{tet}(W)\) and \(\text{tet}(S)\). Out of the 44 Tc\(^r\) fosmid clones from the infant library, 43 were shown to be positive for \(\text{tet}(M)\) by PCR. One of the end-sequences (B04-U-PC1R, 386 bp) from the fosmid clone negative for the \(\text{tet}(M)\) PCR was identical to a region in \(\text{tet}(M)\) downstream of one of the screening primers. Thus all 44 Tc\(^r\) clones from the infant’s metagenomic library were \(\text{tet}(M)\) positive and negative for \(\text{tet}(S)\), \(\text{tet}(O)\) and \(\text{tet}(W)\). In contrast, out of the 272 Tc\(^r\) clones from the mother library, 21 (7.7%) were only positive for \(\text{tet}(W)\) and 204 (75%) were only positive for \(\text{tet}(O)\); for 47 (17.3%) of the Tc\(^r\) clones none of the assayed Tc\(^r\) genes were detected, and all clones were negative for \(\text{tet}(O)\) and \(\text{tet}(S)\).

To further investigate what resistance genes might be present in the 47 maternal clones that were negative for \(\text{tet}(M)\), \(\text{tet}(O)\), \(\text{tet}(W)\) and \(\text{tet}(S)\), we performed a series of multiplex PCRs designed to detect \(\text{tet}(A)\), \(\text{tet}(B)\), \(\text{tet}(C)\), \(\text{tet}(D)\), \(\text{tet}(E)\), \(\text{tet}(G)\), \(\text{tet}(K)\), \(\text{tet}(L)\), \(\text{tet}(P)\), \(\text{tet}(Q)\) and \(\text{erm}(T)\). This second round of PCR screening detected \(\text{tet}(X)\), encoding a tetracycline-inactivating enzyme, in 17 of the tested clones and none of the other genes assayed. Overall, our PCR screens were able to account for the Tc\(^r\) genes present in 242 (89%) of the Tc\(^r\) clones from the mother’s metagenomic library.

Sequencing of all the PCR screening products for \(\text{tet}(M)\) detected in the infant library identified a single sequence type, \(\text{tet}(M)a\), based on 505 bp out of the 1920 bp of the \(\text{tet}(M)\) gene (Table S1). All 21 \(\text{tet}(W)\) PCR screening products, 63 of the 204 products for \(\text{tet}(O)\) and 12 of the 17 \(\text{tet}(X)\) products from the maternal library were also sequenced (Table S2). Of the sequenced \(\text{tet}(O)\) products, 13 were selected to represent genes assigned to different families/genera (see later in the Results section) and the remaining 50 were randomly selected. Based on 609 bp and 499 bp out of the 1920 bp of \(\text{tet}(W)\) and \(\text{tet}(O)\) and 446 bp out of the 1161–1167 bp of \(\text{tet}(X)\), 2 \(\text{tet}(W)a\), b), 9 \(\text{tet}(O)a–i\) and 1 \(\text{tet}(X)a\) different sequence types were identified (Table S2). The sequenced PCR screening products, \(\text{tet}(M)a\) and \(\text{tet}(O)a–i\) could discriminate among the known variants of \(\text{tet}(M)\) and \(\text{tet}(O)\) (Fig. S1, S2). \(\text{tet}(W)a\) and \(\text{tet}(W)b\) could discriminate between groups with highly related \(\text{tet}(W)\) genes sharing 99.9–100% and 99.5–100% sequence identity, respectively (Fig. S3). \(\text{tet}(M)a\) could discriminate among most known variants (Fig. S4) and was identical to the corresponding fragments from two \(\text{tet}(X)\) genes identified in *Bacteroides*, including the \(\text{tet}(X)\) gene first detected in transposon Tn531/Tn4400 [16].

**PCR screening of the infant’s fecal DNA for the presence of \(\text{tet}(W)\), \(\text{tet}(O)\) and \(\text{tet}(X)\)**

Although \(\text{tet}(W)\), \(\text{tet}(O)\) and \(\text{tet}(X)\) were not detected in the infant’s metagenomic library, these genes could nonetheless have been present in the infant GIT microbiome, perhaps in non-abundant species that were not captured in the library. To further investigate the possibility of maternal transmission of resistances to the infant, we PCR-screened the total DNA from the infant fecal sample from which the metagenomic library was constructed. \(\text{tet}(X)\) was not detected, but, remarkably, we obtained amplifications in both the \(\text{tet}(W)\) and the \(\text{tet}(O)\) screening PCRs, albeit the \(\text{tet}(W)\) product could only be observed as a very faint band after a standard number of screening PCR cycles. Both products yielded clean sequence reads, indicating that single sequence types were present in the amplicons. These screening products \(\text{tet}(W)\text{-infant}\_plug\) and \(\text{tet}(O)\text{-infant}\_plug\) were shown to be identical to \(\text{tet}(W)a\) and \(\text{tet}(O)\) detected in the maternal metagenomic library (Fig. S2 and S3). This strongly indicates mother to infant transmission of specific \(\text{tet}(W)\) and \(\text{tet}(O)\) genes.

**Identification of a novel composite Tn916/1545-like conjugative transposon carrying \(\text{tet}(M)\), \(\text{tet}(L)\) and \(\text{erm}(T)\) in the infant library**

In a sequenced fosmid from the infant library, \(\text{tet}(M)\) was found on a Tn916/1545-like transposon, a family of conjugative transposons that have an extremely broad host range [17,18]. The transposon was highly similar to a putative Tn916-like transposon identified in *Streptococcus galiolyticus* subsp. *galiolyticus* strain UCN34 (FN597254), isolated from an elderly endocarditis and colon cancer patient [19], although the infant transposon was located at the 3’end of \(\text{erm}(G)\) (predicted to encode protein L33 from the ribosomal 50S subunit), whereas the transposon from strain UCN34 was located in a putative peptidoglycan-linked protein (Fig. 1). Both transposons contained a second \(\text{tet}(L)\) gene, \(\text{tet}(L)\), predicted to encode an efflux protein, closely linked to plasmid recombination/mobilization \(\text{pre}/\text{mob}\) and replication \(\text{rep}\) genes. A DNA fragment containing the Tn916-like \(\text{orf}12\) as well as \(\text{tet}(M)\), \(\text{tet}(L)\), \(\text{pre}/\text{mob}\) and most of \(\text{rep}\) (see Fig. 1) has also recently been deposited in GenBank (AEELO1000025, contig of 6541 bp) as part of the draft sequence of a Human Microbiome Project (HMP) strain characterized as *S. bovis* ATGC 700338 and isolated from the vagina. This fragment is 100% identical to the homologous region in the infant transposon. Additionally, another DNA fragment from the same *S. bovis* draft sequence (AEELO1000027, contig of 33283 bp) contained a region (1900 bp) with Tn916-like \(\text{orf}5\), \(\text{xis}\) and \(\text{int}\) that was 100% identical with a homologous region in the infant transposon (Fig. 1). This Tn916-like region was located at the 3’end of putative transposase ISStyI. The infant transposon also contained a 3026 bp sequence encoding an erythromycin rRNA methylase gene, \(\text{erm}(T)\),
surrounded by two putative IS1216 transposase genes, not present in strain UCN34. This 3026 bp sequence was 100% identical, except for an additional 30 bp between the erm(T) leader and the second IS1216 element (overall DNA identity of 99%), to a corresponding fragment from S. gallolyticus subsp. pasteurianus NTUH 7421 (AY894138) [20] (Fig. 1). A fragment from S. bovis ATCC 700338 (AEEL01000026, contig of 1578 bp) containing erm(T), leader and an IS1216 was 100% identical with the homologous region in the infant transposon (Fig. 1).

At the genic level, the tet(M) sequence in the infant library insert was also highly similar to previously identified tet(M) genes from the composite transposons, CTn6092, identified in Streptococcus cristus (AY898750, 99.9% DNA identity), and Tn1545, first identified in Streptococcus pneumoniae BM420 (AM638914, 98.2% DNA identity) [21,22]. The infant tet(L) gene was shown to share 98.1% DNA identity (1.4% gaps) with tet(L) genes found on pre/mob- and rep-containing plasmids in Bacillus stearothermophilus (M63091), Enterococcus faecalis (AF503772) and Staphylococcus aureus.

Figure 1. Composite transposon Tn6079 from the infant metagenome compared to corresponding sequences from S. gallolyticus strains. Part of the sequenced fosmid-insert in the infant Tc r metagenome (GU951538; 1–3095) is compared with the most similar homologous sequences in current databases, from S. gallolyticus subsp. gallolyticus strain UCN34 (FN597254; minus strand 564709–589025 and 52545–55152), S. gallolyticus subsp. pasteurianus NTUH7421 (AY894138; minus strand 1–4107) and S. bovis ATCC700338 (AEEL01000024; 112037–123526, AEEL01000025; 1–6541, AEEL01000026; 1–1578, AEEL01000027; 1–3192, AEEL01000028; 37166–39101). The relationships between sequences are shown as percentage identity at nucleotide level, calculated with the EMBOSS program Needle (http://www.ebi.ac.uk/Tools/emboss/align/index.html). Light gray arrows represent ORFs with homology to ORFs from Tn916/Tn1545-like conjugative transposons: orf5–10, orf12–24, tet(M), an excisionase (xis) and an integrase (int) of Tn6079. Dark gray arrows illustrate ORFs that may be of plasmid origin: tet(L), pre/mob and rep. Black arrows illustrate ORFs that appear to be inserted by the two identical IS1216-like elements: first IS1216, erm(T), erm(T) leader sequence and second IS1216.

mpG and mpF were predicted to encode L33 and L32 of the 50S ribosomal subunit. The functions of the predicted ORFs, mid12_1, mid12_5 and mid12_7 are unknown. hisS: histidyl-tRNA synthetase. ISSyd1: putative transposase.

doi:10.1371/journal.pone.0021644.g001
(FM207105), as well as with chromosomally encoded \( \text{tet(L)} \) genes located just downstream of \( \text{tet(M)} \) and upstream of \( \text{pre/mob} \) in \textit{Streptococcus suis} (FM252032) and \textit{S. gallolyticus} subsp. \textit{gallolyticus} strain UCN34 (Fig. 1).

Thus we have identified a novel composite \( \text{Tn916/1545} \)-like conjugative transposon which we registered as \( \text{Tn6079} \) in the Transposon Nomenclature Database from the UCL Eastman Dental Institute, London (http://www.ucl.ac.uk/eastman/tan/) [29].

\( \text{Tc}' \) was conferred exclusively by \( \text{tet(M)} \) and/or \( \text{tet(L)} \) from streptococci in the infant metagenomic library

In the infant \( \text{Tc}' \) metagenomic library, 97.7% (43/44) of the fosmids had at least one of their end-reads assigned within bacteria and at the genus level 72.7% (32/44) of the fosmids were assigned within \textit{Streptococcus} (Fig. 2A and Table 1). This was further supported by a comparison showing 100% DNA identity between the \( \text{rpmG} \) and \( \text{rpmF} \) ORFs from the sequenced infant fosmid, predicted to encode the 50S ribosomal subunit proteins L33 and L32, and the corresponding regions from the three \textit{Streptococcus} strains in Fig. 1 containing similar MGE fragments (S. \textit{gallolyticus} subsp. \textit{pasteurianus} strain NTUH 7421, S. \textit{gallolyticus} subsp. \textit{gallolyticus} strain UCN34 and S. \textit{bosis} ATCC 700336).

Out of the 44 infant fosmids, 43 were positive for a PCR designed to amplify the \( \text{rpmG} \) and \( \text{rpmF} \) ORFs region from the sequenced fosmid (Fig. 1). The reverse primer used in this PCR was specifically designed to target a sequence just downstream of the \( \text{rpmF} \) ORF that was present in the sequenced fosmid and in S. \textit{gallolyticus} subsp. \textit{pasteurianus} strain NTUH 7421 and S. \textit{bosis} ATCC 700336 (AEE101000028) but not in S. \textit{gallolyticus} subsp. \textit{gallolyticus} strain UCN34. The reverse end-sequence from the only fosmid that was negative for this PCR (B04-M32-PCC1R, 625 bp) mapped to a region containing the \text{Tn916}-like \( \text{orf5} \) in the sequenced transposon, which showed that this insert ended within this ORF (see Fig. 1). All 44 fosmids were positive for \( \text{tet(L)} \) and \( \text{erm(T)} \) PCR screenings. Alignment of sequences for 5 randomly selected PCR products containing \( \text{rpmG} \) and \( \text{rpmF} \) showed 100% identity (over 634 bp) with the sequenced fosmid (see Fig. 1). Thus the 44 \( \text{Tc}' \) fosmids from the infant probably represent fragments from the same \textit{S. gallolyticus}-like genomic region having different fragment-specific start and end points.

\( \text{Tc}' \) was conferred mainly by Firmicutes and Bacteroidetes in the mother metagenome

In the adult \( \text{Tc}' \) metagenome, 83.1% (226/272) of the fosmids had at least one of their end-reads assigned within Bacteria (Fig. 2B and Table 2). For 9.9% (27/272) of the fosmids neither of the end-reads had at least one of their end-reads assigned within Bacteria (Fig. 2B). Remarkably, fosmids containing \( \text{tet(W)} \) have been horizontally transferred among different members of the Clostridiales and that specific \( \text{tet(O)} \) and \( \text{tet(X)} \) genes have transferred between bacteria belonging to different phyla.

Discussion

This is a case study based on a functional screen for \( \text{Tc}' \) fosmid clones from two previously prepared metagenomic libraries representing the gut microbiota from an infant and his mother one month after childbirth [9]. In correspondence with the lower complexity in the infant gut microbiota compared to the mother, we detected much fewer \( \text{Tc}' \) fosmid clones in the infant metagenome than in the mother metagenome. The \( \text{Tc}' \) genes detected in the infant metagenomic library did not represent a subset of those found in the mother, but rather a completely distinct set, belonging to a different gene class and encoded by a different species. However, total fecal DNA from the infant sample was shown to also contain specific \( \text{Tc}' \) genes that were present in the maternal library (\( \text{tet(W)a} \) and \( \text{tet(O)h} \)), suggesting that these may have been transmitted from mother to son. Given that this DNA was not cloned, we can not determine the organisms that carried these genes in the infant, but phylogenetic assignment of maternal fosmids suggests that they may have been present in organisms belonging to the Clostridiales (and/or to \textit{Bacteroides} in the case of \( \text{tet(O)} \)).

In the infant library, \( \text{tet(M)} \) and \( \text{tet(L)} \) were detected in all the \( \text{Tc}' \) fosmids whereas mainly \( \text{tet(O)} \) but also \( \text{tet(W)} \) and \( \text{tet(X)} \) were detected among the \( \text{Tc}' \) fosmids from the mother. Although the approach employed here can only detect \( \text{Tc}' \) genes that can be expressed in the \textit{E. coli} library host, the prevalence of \( \text{Tc}' \) genes observed in this study is in general agreement with former culture-independent studies that analyzed \( \text{Tc}' \) in Europe directly by PCR or microarray hybridization [14,15]. These works detected the \( \text{tet(M)} \) genotype to be abundant in Finnish infant fecal samples [15] and \( \text{tet(O)} \) and \( \text{tet(W)} \) to be the most prevalent \( \text{Tc}' \) genes in fecal samples from adults in six different European countries as detected by microarray analysis [14]. In contrast, a functional metagenomic screen of antibiotic resistances in the gut of two adult individuals carried out in the USA (Boston, MA) recovered numerous \( \text{tet(W)} \) sequences but did not identify any \( \text{tet(O)} \) or \( \text{tet(X)} \) genes [24]. It is important to note that the screening approaches in this latter work and in our own study both require that \( \text{Tc}' \) can be expressed in the \textit{E. coli} library host strains at a level sufficient to confer resistance in the presence of the antibiotic, although each employs a different cloning vector. Our fosmid-based study has the potential disadvantage that resistance genes located on smaller plasmids (<40 kb) may not be represented in the metagenomic libraries, but, on the other hand, the larger insert size increases the likelihood to clone complete resistance genes and enables the recovery of complex genetic elements. The different results obtained in the two USA studies could be due to the different
Figure 2. Microbial diversity of Tc fosmid clones in infant (A) and mother (B) metagenomes. Modified MEGAN tree (collapsed at Genus level) showing summarized number of reads assigned at different taxonomical levels. The size of each node is proportional to the number of reads assigned to the node. Beneath each node the number and percentage of Tc genes detected in this study are noted. A. The "No hits & removed MGE hits" category contains 13 reads with no BLASTX hits (or hits that did not attain the min score/length of 0.15), 2 removed reads which were predicted to be located in MGE and were initially assigned below order level and 4 reads that mapped to ORFs in the sequenced transposon (Tn6079). B. The "No hits & removed MGE hits" category contains 138 reads with no BLASTX hits (or hits that did not attain the min score/length of 0.15) and 31 removed reads which were predicted to be located in MGE and were initially assigned below order level. The "Not assigned" category contains 2 reads that were assigned by BLASTX hits to uncultured bacteria.

doi:10.1371/journal.pone.0021644.g002
cloning systems and/or reflect the antibiotic concentrations used in the functional screenings (10 μg/ml tetracycline in our study versus 20 μg/ml tetracycline, oxytetracycline or minocycline in Sommer et al.). In addition, functional screenings performed at even lower tetracycline concentrations might reveal further Tc' genes that are weakly expressed.

The previous culture-independent analyses that identified Tc' in human fecal samples did not investigate the bacterial species in which such resistance was encoded [14,15,24]. In our fosmid-based study, end-sequencing of fosmid inserts allowed for taxonomic identification of the resistant organisms present in the GIT of the two individuals analyzed. All infant Tc' clones appeared to represent the same Streptococcus genomic region containing tet(M), tet(L) and erm(T) within a novel composite Tn916-like transposon, Tn6079, located at the 5’ end of rpmG. The nucleotide sequences of both rpmG and its 5’ neighbor rpmF were 100% identical to those of S. gallolyticus subsp. pasteurianus strain NTUH 7421, S. gallolyticus subsp. galloyticus strain UCN34 and S. bovis ATCC 700338 (Fig. 1). It is important to note that the heterogenous group of strains traditionally designated S. bovis has recently been split by modern taxonomic techniques into the sister species S. galloyticus and S. infantarius [25]. Indeed, the 16S rRNA sequence of S. bovis ATCC 700338 shows 99 to 100% identity with the S. galloyticus subsp. pasteurianus and S. galloyticus subsp. galloyticus 16S rRNA sequences currently available in GenBank. Therefore, the presence of rpmG and rpmF next to the Tn6079 transposon in the sequenced fosmid allows for identification of the Tc'-carrying organism in the infant GIT to species level. The similarities in sequence and structure between Tn6079 and corresponding MGE sequences in the S. galloyticus subsp. pasteurianus, S. galloyticus subsp. galloyticus and S. bovis ATCC 700338 strains (Fig. 1) strongly suggest that the infant’s composite transposon arose through a process involving intraspecific genetic exchange.

Regarding the origin of the S. galloyticus-like strain carrying the transposon in the infant, this organism was probably not transmitted from the maternal GIT, since no streptococci were detected in the mother’s fecal samples, neither in the resistance screens performed here, nor in the previous random end sequencing of the library [9], even though this species is a normal inhabitant in the GIT of humans and animals and can be isolated in 5–16% of fecal samples from healthy adults [26]. Possible origins may include transmission from other maternal areas that are known to often harbor streptococci, such as the skin, the birth canal and the mouth, from breast milk, where streptococci have also recently been detected [27], or from other individuals handling the infant. The 100% identities recovered between the sequenced fosmid insert and the vaginal strain S. bovis ATCC 700338 (see Fig. 1) suggest that the infant may have acquired this strain or a closely related one during his passage through the birth canal.

In the maternal library, microorganisms conferring Tc' consisted mainly of Firmicutes and Bacteroidetes, which commonly represent the two major Phyla of the human GIT [28,29] and were also the most represented in the fosmid library of the mother according to random end reads [9]. For 9.9% of the maternal Tc' fosmids, neither of the end-reads had any BLASTX hits against the NCBI non-redundant protein database in spite of being of substantial length (>500 bp). These fosmids likely carry Tc' genes, mainly tet(O), from microorganisms for which no close relatives have yet been cultured. tet(O) was the main gene conferring resistance and was detected both within the Clostridiales (Firmicutes) and also in the Bacteroides (Bacteroidetes), where it had not been reported previously. Similarly, tet(X) was detected within the Clostridiales and the Bacteroides and fosmids carrying identical tet(O) or tet(X) sequences were assigned within both phyla/orders. (Fig. 2) shows the presence of gene sequences among the three different families (Table S2). Sequence identity is not expected between genes that have been diverging as orthologs since the phylogenetic split between such distantly related bacteria and can therefore be interpreted as evidence of recent horizontal transfers among these organisms [30–32]. tet(W) genes and flanking sequences in different isolates of GIT bacteria from diverse hosts have also been shown to share a high degree of similarity in previous analyses [33]. However, this is the first time that exact sequences of an antibiotic resistance gene are shown to occur in distantly related bacteria naturally.

### Table 1. Assignment of end-reads and corresponding fosmids from the infant Tc' metagenome at different taxonomical levels.

<table>
<thead>
<tr>
<th>Level of assignment</th>
<th>Assignment details</th>
<th>No. of reads (percentage of total reads: 88)</th>
<th>No. of fosmids (percentage of total fosmids: 44)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kingdom</td>
<td>Reads/fosmids of which at least one end-read was assigned within Bacteria.</td>
<td>69 (78.4%)</td>
<td>43 (97.7%)</td>
</tr>
<tr>
<td>Phylum</td>
<td>Reads/fosmids of which at least one end-read was assigned within Firmicutes.</td>
<td>55 (62.5%)</td>
<td>39 (88.6%)</td>
</tr>
<tr>
<td>Order</td>
<td>Reads/fosmids of which at least one end-read was assigned within Lactobacillales.</td>
<td>38 (43.2%)</td>
<td>36 (81.8%)</td>
</tr>
<tr>
<td>Family</td>
<td>Reads/fosmids of which at least one end-read was assigned within Streptococcaceae.</td>
<td>38 (43.2%)</td>
<td>36 (81.8%)</td>
</tr>
<tr>
<td>Genus</td>
<td>Reads/fosmids of which at least one end-read was assigned within Streptococcus.</td>
<td>34 (38.6%)</td>
<td>32 (72.7%)</td>
</tr>
<tr>
<td>No hit &amp; removed MGE hits</td>
<td>Both forward and reverse fosmid end-reads in “No hit &amp; removed MGE hits”</td>
<td>2 (2.3%)</td>
<td>1 (2.3%)</td>
</tr>
<tr>
<td></td>
<td>Both forward and reverse fosmid end-reads with no blasts hits or hits below min cut off (0.13 bit score/length).</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Both forward and reverse fosmid end-reads with no blasts hits.</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

doi:10.1371/journal.pone.0021644.t001
coexisting in the gut of a single person at a particular point in time. Although their coexistence does not prove that the horizontal transfers occurred in the GIT of the infant’s mother, alternative explanations would still necessitate recent transfers, in the environment or in the GITs of other individuals, followed by colocalization of the bacteria in this individual’s GIT. These scenarios would imply a high frequency of these exact sequences in nature and/or a high likelihood of colocalization of the bacteria carrying them, and therefore seem less parsimonious than in situ transfer among bacteria coexisting closely in the dense microbiota of the adult GIT. In addition to fosmids containing tet(O), tet(W) and tet(X), there was also a fraction of Tc’ maternal fosmids in which none of the assayed genes were detected (11%, assigned within Firmicutes or Bacteroidetes), and where Tc’ must have been conferred by rare resistance genes.

This study showed strong indications of transmission of specific Tc’ genes (tet(W)a and tet(O)h) from the mother’s GIT to that of the infant. However, the third Tc’ gene present in the maternal genomic library, tet(X), was not detected in the infant. tet(W)a and tet(O)h could be found in the infant’s uncloned fecal DNA but not in the infant metagenomic library, suggesting that they were only present in low numbers. This is supported by the fact that PCR with tet(W) screening primers produced only a very faint band, and by the detection of a single tet(O) sequence type in the infant out of the 9 different types detected in the mother (Fig. S2). The scarcity of maternal Tc’ genes in the infant could be partially explained by the fact that approximately half of the detected Tc’ in the mother library was encoded by clostridia, and previous analyses of random end sequences from these libraries showed that clostridia were not abundant in the mother and that they were not transmitted to the infant [9]. On the other hand, those analyses, as well as comparisons of fosmid sequences and Bacteroides-specific 16S PCR libraries, have shown transmission to the infant of the two Bacteroides phylotypes present in the mother [9].

Table 2. Assignment of end-reads and corresponding fosmids from the mother Tc’ metagenome at different taxonomical levels.

<table>
<thead>
<tr>
<th>Level of assignment</th>
<th>Assignment details</th>
<th>No. of reads (percentage of total reads: 544)</th>
<th>No. of fosmids (percentage of total fosmids: 272)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kingdom</td>
<td>Reads/fosmids of which at least one end-read was assigned within Bacteria.</td>
<td>365 (67.1%)</td>
<td>226 (83.1%)</td>
</tr>
<tr>
<td>Phylum</td>
<td>Reads/fosmids of which at least one end-read was assigned within Bacteroidetes.</td>
<td>27 (5.0%)</td>
<td>22 (8.1%)</td>
</tr>
<tr>
<td></td>
<td>Reads/fosmids of which at least one end-read was assigned within Firmicutes.</td>
<td>254 (46.7%)</td>
<td>180 (66.2%)</td>
</tr>
<tr>
<td>Order</td>
<td>Reads/fosmids of which at least one end-read was assigned within Bacteroidales.</td>
<td>20 (3.7%)</td>
<td>19 (7.0%)</td>
</tr>
<tr>
<td></td>
<td>Reads/fosmids of which at least one end-read was assigned within Clostridales.</td>
<td>182 (33.4%)</td>
<td>139 (51.1%)</td>
</tr>
<tr>
<td>Family</td>
<td>Reads/fosmids of which at least one end-read was assigned within Bacteroidaceae.</td>
<td>9 (1.7%)</td>
<td>9 (3.3%)</td>
</tr>
<tr>
<td></td>
<td>Reads/fosmids of which at least one end-read was assigned within Lachnospiraceae.</td>
<td>35 (6.4%)</td>
<td>29 (10.7%)</td>
</tr>
<tr>
<td></td>
<td>Reads/fosmids of which at least one end-read was assigned within Clostridaceae.</td>
<td>9 (1.7%)</td>
<td>9 (3.3%)</td>
</tr>
<tr>
<td></td>
<td>Reads/fosmids of which at least one end-read was assigned within Eubacteriaceae.</td>
<td>1 (0.2%)</td>
<td>1 (0.4%)</td>
</tr>
<tr>
<td></td>
<td>Reads/fosmids of which at least one end-read was assigned within Ruminococcaceae.</td>
<td>10 (1.8%)</td>
<td>10 (3.7%)</td>
</tr>
<tr>
<td>Genus</td>
<td>Reads/fosmids of which at least one end-read was assigned within Bacteroides.</td>
<td>9 (1.7%)</td>
<td>9 (3.3%)</td>
</tr>
<tr>
<td></td>
<td>Reads/fosmids of which at least one end-read was assigned within Dorea.</td>
<td>16 (2.9%)</td>
<td>12 (4.4%)</td>
</tr>
<tr>
<td></td>
<td>Reads/fosmids of which at least one end-read was assigned within Clostridium.</td>
<td>9 (1.7%)</td>
<td>9 (3.3%)</td>
</tr>
<tr>
<td></td>
<td>Reads/fosmids of which at least one end-read was assigned within Eubacterium.</td>
<td>1 (0.2%)</td>
<td>1 (0.4%)</td>
</tr>
<tr>
<td></td>
<td>Reads/fosmids of which at least one end-read was assigned within Ruminococcus.</td>
<td>8 (1.5%)</td>
<td>8 (2.9%)</td>
</tr>
<tr>
<td></td>
<td>Reads/fosmids of which at least one end-read was assigned within Faecalibacterium.</td>
<td>2 (0.4%)</td>
<td>2 (0.7%)</td>
</tr>
<tr>
<td>“No hit &amp; removed MGE hits”</td>
<td>Both forward and reverse fosmid end-reads in “No hit &amp; removed MGE hits”</td>
<td>92 (16.9%)</td>
<td>46 (16.9%)</td>
</tr>
<tr>
<td></td>
<td>Both forward and reverse fosmid end-reads with no blastx hits or hits below min cut off (0.15 bit score/length).</td>
<td>76 (13.9%)</td>
<td>38 (13.9%)</td>
</tr>
<tr>
<td></td>
<td>Both forward and reverse fosmid end-reads with no blastx hits.</td>
<td>54 (9.9%)</td>
<td>27 (9.9%)</td>
</tr>
</tbody>
</table>

doi:10.1371/journal.pone.0021644.t002

Tc’ in Human Gut Microbiota

PLoS ONE | www.plosone.org 7 June 2011 | Volume 6 | Issue 6 | e21644
established phylotype transmission, the fact that *Bacteroides* were amply represented in both the mother and infant libraries and the presence of tet(O)h and tetXa genes in *Bacteroides*-assigned maternal fosmid sequences, the lack of Tc'-encoding fosmids would be expected in a 69,000-clone library if tet(O)h and/or tetXa were present in every *Bacteroides* cell (based on a genome size of 6.5 Mb); in contrast, only 3 and 18 end-reads were assigned to Bacteroidetes among the mother clones containing tet(O)h and tetXa, respectively (Fig. 2B and Table S2).

In summary, for the first time we have characterized the microbial diversity of Tc' bacteria in human gut samples, by analyzing GIT fosmid libraries from a mother and her infant. The maternal and infant libraries contained different resistant taxa encoding distinct sets of genes, but some of the specific Tc' genes present in the mother could be recovered from uncloned infant fecal DNA. This indicates that transmission of Tc' genes from the mother's GIT to the infant likely occurred, but that, due to the complexity of the GIT microbiota, species and genes present in low numbers were missed in the infant metagenomic library in spite of its large size (>70,000 clones). The likely role of the human gut as a privileged environment for HGT has been previously recognized [8], but here we present the first documented cases of Tc' transfer from the mother's GIT to the infant likely occurred, but that, due to the dereplication of transposon Tn6079, the 44 infant clones were also screened for tet(L) [36], for a region linking erm(T) to an IS element (1010 bp) and for a region containing the rpmG and rpmF ORFs (664 bp) using primer pairs TsrL-1/TsrL-2, ermG-2/IS1216V3-1 and <rro/rrNA_S (Table S3). The latter PCR primer pair was designed to specifically target the region starting just downstream of the int gene within Tn6079 and ending downstream of the rpmF gene.

PCR screening products from both uncloned fecal DNA from the infant (tet(W) and tet(O)) and from each metagenomic library (tet(M), tet(W), tet(O) and tet(X)) were PCR amplified with the primer pairs described above and sequenced by Macrogen, Korea (http://www.macrogen.com/eng/sequencing/sequence_main.jsp) (see Tables S1 and S2). In addition, 5 randomly selected PCR screening products containing the rpmG and rpmF ORFs (B04-M4, B04-M8, B04-M13, B04-M16 and B04-M18) were also sequenced. ClustalX [37] was used to align sequences within the tet groups to determine different sequence types (Tables S1 & S2) and to align the rpmG and rpmF sequences. All together, 13 different Tc' gene sequence types were detected in the mother (tet(W), tet(O) and tet(X)).

**Materials and Methods**

**Sample collection and ethics statement**

The infant and mother metagenomic fosmid libraries analyzed in this study were prepared from fecal samples obtained one month after the infant's birth [9]. The infant was a healthy male, vaginally delivered at full term at the University Medical Center of the University of Arizona in Tucson (USA). He was exclusively breast-fed for 5 months. Samples were collected at the University of Arizona, with informed written consent from the infant’s parents, using protocols approved by the institutional review boards of the Lawrence Berkeley National Laboratory and the University of Arizona.

**Metagenomic fosmid libraries and preparation of master plates with pooled clones**

The infant and mother metagenomic fosmid libraries analyzed consisted of approximately 76000 and 69100 clones, respectively [9]. Fosmid inserts were approximately 40 kb, thus the infant and mother libraries represent roughly 3 Gb of DNA each. Clones from the infant and mother metagenomic fosmid libraries were pooled resulting in a reduction from 198 and 180 library (384 wells) plates to 14 and 12 (384 wells) master plates, respectively. Each master plate was constructed by pooling 15 library plates into one master plate using a Plate Mate Plus from Matrix. Each well in the master plates contained 30–40 µl LB (Miller’s) broth supplied with 7.5% Glycerol and 2 µl from each of the original library plates.

**Phenotypic screening of library master plates for Tc'**

All master plates were screened for clones conferring Tc' resistance on growth plates (384 wells) containing 60 µl LB (Miller’s) broth with 10 µg/ml tetracycline per well. Growth or no growth was detected after overnight incubation at 37°C. When growth was detected, each of the 15 clones from the original fosmid libraries that could be responsible for the observed resistance phenotype was tested for Tc' separately as described above.

**Genotypic PCR screenings and sequencing**

All Tc' clones (44 infant and 272 mother clones) from the two metagenomic libraries were screened for tet(M), tet(S), tet(O) and tet(W) by PCR using primer pairs TetM-1D/TetM-2, tetW-1/tetW-2, TetS-1/TetS-2 and TetO-1/TetO-2 (Table S3) [34]. To detect other possible resistance genes in the maternal Tc' clones that were negative for these screening primers, we performed a series of multiplex PCRs designed to detect tet(A), tet(B), tet(C), tet(D), tet(E), tet(G), tet(K), tet(L), tet(NP), tet(Q) and tet(X), using previously reported primer combinations and protocols [35]. In addition, purified DNA from the infant fecal sample from which the infant fosmid library was constructed was screened for the Tc' genes detected in the mother (tet(W), tet(O) and tet(X)).

After characterization of transposon Tn6079, the 44 infant clones that were negative for these screening primers were performed a series of multiplex PCRs designed to detect tet(A), tet(B), tet(C), tet(D), tet(E), tet(G), tet(K), tet(L), tet(NP), tet(Q) and tet(X), using previously reported primer combinations and protocols [35]. In addition, purified DNA from the infant fecal sample from which the infant fosmid library was constructed was screened for the Tc' genes detected in the mother (tet(W), tet(O) and tet(X)).

After characterization of transposon Tn6079, the 44 infant clones were also screened for tet(L) [36], for a region linking erm(T) to an IS element (1010 bp) and for a region containing the rpmG and rpmF ORFs (664 bp) using primer pairs TsrL-1/TsrL-2, ermG-2/IS1216V3-1 and <rro/rrNA_S (Table S3). The latter PCR primer pair was designed to specifically target the region starting just downstream of the int gene within Tn6079 and ending downstream of the rpmF gene.

PCR screening products from both uncloned fecal DNA from the infant (tet(W) and tet(O)) and from each metagenomic library (tet(M), tet(W), tet(O) and tet(X)) were sequenced with the PCR primers by Macrogen, Korea (http://www.macrogen.com/eng/sequencing/sequence_main.jsp) (see Tables S1 and S2). In addition, 5 randomly selected PCR screening products containing the rpmG and rpmF ORFs (B04-M4, B04-M8, B04-M13, B04-M16 and B04-M18) were also sequenced. ClustalX [37] was used to align sequences within the tet groups to determine different sequence types (Tables S1 & S2) and to align the rpmG and rpmF sequences. All together, 13 different Tc' gene sequence types were detected in the mother (tet(W), tet(O) and tet(X)).

**Sequencing a fosmid-insert carrying tet(M)**

One fosmid carrying tet(M) from the infant library (B04-M2) was sequenced as part of a mix of 12 fosmids pyrosequenced with Multiple Sequence Identifiers (MIDs) in a Roche GS FLX
instrument in the Sequencing Technology group of the DOE Joint Genome Institute (JGI), CA, USA [http://www.jgi.doe.gov/]. Reads belonging to the B04-M2 fosmid were sorted out and assembled using the Roche 454 Newbler software. Seven contigs were generated of which five (lengths 23744 bp, 19061 bp, 6382 bp, 1578 bp and 809 bp) were used to assemble the fosmid sequence. The remaining two contigs (lengths 938 bp and 507 bp) were highly similar or identical to the E. coli host genome and therefore were not incorporated into the assembly. Ten sequencing primers, M1b, M2, M2b, M3, M4, M5, M6, C340F, C01F and C01R were designed (Table S3) and Sanger reads, produced by Macrogen, Korea, were used to close the remaining five gaps. FosmidMAX™ DNA Purification Kit (EPICENTRE, USA) was used to prepare fosmid DNA template for the Sanger sequencing reactions. A finished 53499 bp circular fosmid containing a 45066 bp insert was assembled. The insert was annotated by Macrogen, Korea, and quality checking (Q20) and trimming were performed manually in Vector NTI. Vector contaminations were removed from 16 end-reads prior to Genome Survey Sequences (GSS) submission to GenBank. All together 632 high quality sequences were used as queries to prepare results very similar to those obtained based on the analysis of 16S sequences [41]. Each of the two BLASTX results (mother and infant) were separately parsed and visualized using MEGAN 

Assignment of end-sequences

All together 632 high quality sequences were used as queries to establish bacterial diversity through BLASTX searches against the NCBI non-redundant protein database (e-value < e^{-15}). End-sequencing has recently been validated as a reliable method of determining diversity in a metagenomic sample, as random sequence reads from fosmid libraries of human fecal samples provide results very similar to those obtained based on the analysis of 16S sequences [41]. Each of the two BLASTX results (mother and infant) were separately parsed and visualized using MEGAN (version 3.7.4) software (Min Score = 35, Min Score/Length = 0.15, Top Percent = 20, Min Support = 1) [42]. Min Score/Length = 0.15 was chosen in order to account for the different read lengths. Because end-sequences located in MGE could easily bias the bacterial assignment by MEGAN, the BLASTX results were parsed for reads with hits containing the regular expressions conjugative, transposon, in916, integrase, recombinate, excisionase, mobilization and resistance, and if such reads were assigned below order-level they were manually removed (Tables S4 & S5). Additionally, 4 infant end-reads (B04-M19-PCC1F, B04-M20-PCC1F, B04-M29-PCC1R, B04-M33-PCC1R) that were not found by the parsing of the BLASTX result but mapped to ORFs in the sequenced transposon were also removed. The assignment of reads by MEGAN based only on one BLASTX hit is very sensitive to misclassified sequences in GenBank. Therefore the taxonomical classification of BLASTX hit sequences used by MEGAN to assign reads at species level was reviewed (Tables S6 & S7). Finally, it was checked that assignments of forward and reverse end-reads from the same fosmid did not contradict each other.

Supporting Information

Figure S1 NJ tree based on 505 bp corresponding to the sequenced PCR screening products of tet(M). The tree includes 57 tet(M) genes from GenBank and sequence type tet(M)a (bold) found among Tc′ clones in the infant metagenomic library. tet(M)a differs from the 57 tet(M) genes present in GenBank at the time of screening. (TIF)

Figure S2 NJ tet(O) trees showing that tet(O)a–i can discriminate among the known variants of tet(O). A. Tree based on 499 bp corresponding to the sequenced PCR screening products of tet(O). tet(O)a–i represent the nine sequence types found among 63/204 tet(O) fosmids from the maternal metagenomic library and tet(O)infant plu represents the sequence type detected directly in uncloned DNA from the infant fecal sample. B. Tree based on the total tet(O) gene (1920 bp) of 18 GenBank sequences defined as tet(O) by sharing ≥80% identity at the amino acid level. However, NC_006134 is a mosaic combination of tet(O) and tet(M). (TIF)

Figure S3 NJ trees showing to what degree tet(W)a,b can discriminate among the known variants of tet(W). A. Tree based on 609 bp corresponding to the sequenced PCR screening products of tet(W). tet(W)a,b represent the two sequence types found among 21 tet(W) fosmids from the maternal metagenomic library and tet(W)infant plu represents the sequence type detected directly in uncloned DNA from the infant fecal sample. (DQ525023 is not included in group tet(W) because the tet(W) screening primers are not specific for this gene). B. Tree based on the total tet(W) gene (1920 bp) of 24 GenBank sequences defined as tet(W) by sharing ≥80% identity at the amino acid level. However, AY485122, AY485126, AY196920, AY196921, and DQ525023 are different mosaic combinations of tet(W), tet(O) and tet(32) and the tet(W) screening primers are not specific for these genes. (TIF)

Table S1 Sequence type of 43 tet(M) genes detected in the infant metagenome.

(DOCX)

Table S2 Sequence types among 21 tet(W) and 63 out of a total of 204 tet(O) genes detected in the mother metagenome.

(DOCX)

Table S3 Primers used in this study.

(DOCX)

Table S4 End-reads (17) from the infant Tc′ metagenome for which BLASTX hits contained the regular expressions conjugative, transposon, in916, integrase, recombinate, excisionase, mobilization and resistance

(DOCX)
and therefore were predicted to be located in MGE (2 end-reads in bold letters were assigned below order level and therefore removed from their initial assignments to the group with no hits in figure 2A).

**Table S5** End-reads (141) from the mother Te+ metagenome for which BLASTX hits contained the regular expressions conjugative, transposon, tn916, integrase, recombinase, excisionase, mobilization and resistance and therefore were predicted to be located in MGE (31 end-reads in bold letters were assigned below order level and therefore removed from their initial assignments to the group with no hits in figure 2B).

**Table S6** Review of taxonomical classification of BLASTX hit sequences used by MEGAN to assign reads at species level in the infant Te+ metagenome.

**References**


