www.sciencemag.org/cgi/content/full/325/5944/1128/DC1

Supporting Online Material for

Functional Characterization of the Antibiotic Resistance Reservoir in the Human Microflora

Morten O. A. Sommer,* Gautam Dantas,* George M. Church

*To whom correspondence should be addressed.
E-mail: sommer@genetics.med.harvard.edu (M.O.A.S); dantas@wustl.edu (G.D.)

Published 28 August 2009, *Science* **325**, 1128 (2009)
DOI: 10.1126/science.1176950

This PDF file includes:

Functional characterization of the antibiotic resistance reservoir in the human microflora

Supporting Online Materials

Morten O.A. Sommer*^, Gautam Dantas*^ & George M. Church^

Department of Genetics, Harvard Medical School, Boston, MA-02115, USA.

* These authors contributed equally to this work

^Correspondence and requests for materials should be addressed to M.O.A.S (sommer@genetics.med.harvard.edu) or G.D (dantas@wustl.edu).

Materials and Methods

Isolation of metagenomic DNA from human microbiome samples.

Saliva and fecal samples were obtained from two healthy unrelated volunteers according to protocols approved by the Harvard Medical School Institutional Review Board (Study # M15656-101).

DNA from gut microbiomes was extracted directly from fecal samples using the PowerMax Soil DNA Isolation Kit (Mobio Laboratories Inc.) according to their suggested protocol (http://www.mobio.com/files/protocol/12988.pdf), using 5 g of fecal sample instead of the soil sample referred to in the protocol.

DNA from oral microbiomes was extracted directly from saliva samples using the MasterPure DNA Purification kit from Epicentre according to the following protocol:

1. Dilute 1 µl of 50 mg/ml Proteinase K into 150 µl of 2X T and C Lysis Solution.
2. Pellet 1000 µL of saliva sample
3. Remove supernatant and add 150 µl of 2X T and C Lysis buffer with Proteinase K and mix thoroughly.
4. Incubate at 65 °C for 15 minutes; vortex mix every 5 minutes.
5. Cool the samples to 37 °C and add 1 µl of 5 mg/ml RNase A to the sample; mix thoroughly.
6. Incubate at 37 °C for 30 minutes.
7. Place the samples on ice for 3-5 minutes
8. Add 250 µl of MPC Protein Precipitation Reagent to cold lysed sample and vortex mix vigorously for 10 seconds.
9. Pellet the debris by centrifugation for 10 minutes at 10,000 x g in a microcentrifuge.
10. Transfer the supernatant to a clean microcentrifuge tube and discard the pellet.
11. Add 750 µl of isopropanol to the recovered supernatant. Invert the tube several (30-40) times.
12. Pellet the DNA by centrifugation at 4 °C for 10 minutes in a microcentrifuge.
13. Carefully pour off the isopropanol without dislodging the DNA pellet.

14. Rinse twice with 75% ethanol
15. Remove all of the residual ethanol with a pipette.
16. Resuspend the DNA in 20 µl of 10 mM Tris pH 8

Clonal isolation of microbiome isolates and growth profiling in the presence of antibiotics.

572 microbiome isolates capable of growing on Luria broth (LB) agar under aerobic conditions were isolated from fecal samples collected from two unrelated healthy volunteers at three different sampling times (day 1, 140 and 141). Dilutions of fecal samples were plated out on LB agar within 3 hours of sample collection, and incubated for 24 hrs at 37 °C. To profile the ability of these isolates to grow in the presence of antibiotics, individual colonies were picked and inoculated into LB liquid medium for 24 hrs at 37 °C. 5 µL of these cultures were inoculated into 200 µL LB medium containing one of 13 antibiotics at concentrations determined to prevent the growth of wild type *E. coli* (Table S1). Cultures were incubated for 24 hrs at 37 °C and growth was assayed by absorption at 600 nm using a Versamax plate reader from Molecular Devices. For each isolate, absorbance values corresponding to growth in each of the 13 antibiotics was background subtracted (absorbance of un-inoculated LB media) and normalized to corresponding growth in the absence of antibiotic. We expected some temporal variation in the fecal microbiome samples resulting from variables that were not assayed in the study (e.g. diet). In an attempt to achieve a consensus view of the antibiotic resistance profiles of the healthy individuals sampled, we averaged profiles over three sampling days (Day 1, Day 140,

Day 141) (Figure 2). The ability of the isolates to grow in the presence of antibiotics is shown for each sampling time in Figure S2 and S3.

Phylogenetic profiling.

The 16S ribosomal DNA (rDNA) of cultured isolates was amplified using universal bacterial 16S primers:

>Bact_63f_62C
5' – CAG GCC TAA CAC ATG CAA GTC – 3'

>Bact_1389r_63C
5' – ACG GGC GGT GTG TAC AAG – 3'

16S rDNA amplicons were sequenced for phylogenetic profiling. Phylogeny was determined using SeqMatch from the Ribosomal Database Project II server(*1*) and is displayed in Figure S1.

Isolation of genomic DNA from human microbiome isolates.

Microbiome isolates from gut microbiome samples were grown up in 200 µL LB and incubated for 24 hrs at 37 °C. All saturated cultures derived from a particular fecal sample were pooled and cells were harvested by centrifugation. DNA was extracted using the PowerMax Soil DNA Isolation Kit (Mobio Laboratories Inc.) according to their suggested protocol (http://www.mobio.com/files/protocol/12988.pdf) using the pelleted cells from the combined microbiome isolates of each microbiome sample instead of the soil sample referred to in the protocol.

Library construction

All DNA used for functional selections was sheared using a custom made semi automated DNA shearer to a size range of 500 – 3500 base pairs. Sheared DNA was end-repaired using the End-It end repair kit (Epicentre) with the following protocol:

1. Combine:
    a. 276 µL of sheared DNA (30 ng/µL)
    b. 40 µL 10 X End Repair Buffer
    c. 40 µL 2.5 mM dNTP mix
    d. 40 µL 10 mM ATP
    e. 4 µL of End-It enzyme mix
2. Incubate at room temp for 45 minutes
3. Heat inactivate at 70 °C for 20 minutes

End-Repaired DNA was size selected by electrophoresis through a 1 % low melting point agarose gel in 0.5X TBE buffer. A gel slice corresponding to 1000-3000 base pairs was excised from the gel and DNA was extracted using a QIAquick Gel Extraction Kit (Qiagen).

DNA was ligated into pZE21 MCS 1 vector (*2*) at the HincII site using the Fast Link ligation kit (Epicentre) with the following ligation protocol:

1. Combine:
    a. 0.5 µL 10X Ligation buffer
    b. 0.25 µL 10 mM ATP
    c. 0.25 µL dH20
    d. 0.50 µL HincII cut pZE21 MCS 1 vector (300 ng/µL)
    e. 2.50 µL Sheared, End-repaired, Gel purified DNA insert (300 ng/µL)
    f. 1.0 µL Fast-Link ligase (2 units/µL)
2. Incubate at room temperature for 16 hrs
3. Heat inactivate at 70 °C for 20 minutes

4 µL of the fresh ligation mixture was used for transformation by electroporation into 50 µL of electrocompetent *E. coli* TOP10 cells (Invitrogen). After transformation using standard protocols for a 1 mm electroporation cuvette, cells were recovered in 1 mL SOC medium for 1 hr at 37 °C. Libraries were titered by plating out 1 µL and 0.01 µL of recovered cells onto LB agar plates containing 50 µg/mL kanamycin. For each library, insert size distribution was estimated by gel electrophoresis of PCR products obtained by amplifying the insert using primers flanking the HincII site of the multiple cloning site of the pZE21 MCS1 vector (which contains a selectable marker for kanamycin resistance). The average insert size for all libraries was found to be 1.8 kilo base pairs. The total size of each of the genomic and metagenomic libraries as determined by multiplying average PCR based insert size with the number of colony forming units (cfu) in a

given library varied between 0.3 – 5 x 10^9 base pairs. The rest of the recovered cells were inoculated into 10 mLs of LB containing 50 µg/mL kanamycin and grown overnight. The overnight culture was frozen down with 15 % glycerol and kept at -80 °C for subsequent screening.

Functional selections of antibiotic resistant clones from metagenomic and genomic libraries derived from human microbiomes

For each metagenomic and genomic library tested, 100 µL of library freezer stock corresponding to 0.5-1 x 10^8 cfu were plated out on LB agar plates containing binary combinations of kanamycin (50 µg/mL) and one of 13 different antibiotics (Table S1) and incubated at 37 °C for 16 hrs.

To enable functional selections for resistance to multiple compounds, the transformed cells were allowed to grow up overnight before freezing stocks, allowing each clone harboring a unique DNA insert to grow, resulting in an amplification of the particular clone in the library. Based on the original library titers as well as titers of the freezer stocks, the average amplification of a given library can be estimated. On average each unique clone in the libraries screened was plated out in 10-100 copies depending on library size, corresponding to a 10-100 fold amplification. To minimize the redundancy of clones in subsequent analysis, on average the number of colonies picked for sequencing corresponded to approximately the number of resistant clones on an agar plate divided by the estimated average library amplification.

Each of the clones picked was inoculated into liquid LB medium containing kanamycin (50 µg/mL) and the relevant antibiotic to which resistance had been selected (Table S1) and grown overnight to verify resistance phenotype.

Sequencing and analysis of metagenomic and genomic inserts from the human microbiome

Selected metagenomic and genomic inserts harbored by clones identified in the functional selections were sequenced using Sanger sequencing. A total of 634 clones (Table S5) were sequenced bi directionally using the following primers:

>pZE21_81_104_57C

5' – GAA TTC ATT AAA GAG GAG AAA GGT - 3'

>pZE21_151_174rc_58C

5' – TTT CGT TTT ATT TGA TGC CTC TAG – 3'

For each read, sequence corresponding to the cloning vector was removed using DNAStar Laser Gene v. 8.0 – SeqMaster. Assembly of each pair of reads from each clone was attempted using SeqMaster. The remaining reads that were not fully assembled were compared to each other using BLAST and were considered to be derived from two colonies containing the same insert if they were more than 99% identical at the nucleotide level over a stretch longer than 400 base pairs. Only one clone from each set of clones, containing inserts defined to be identical based on these criteria, was selected for subsequent sequencing. For each unique clone, primers were designed for each read to close the insert sequence. This process was repeated once more to close inserts that were particularly large. Finally, 210 unique full length inserts were assembled.

Open reading frames were identified and annotated using ORFfinder (http://www.ncbi.nlm.nih.gov/projects/gorf/). Annotated resistance genes were compared to the GenBank non redundant nucleotide database (November 12, 2008) using tblastx, which computes local sequence alignment between the nucleotide query translated in all 6 frames and the non redundant nucleotide database translated in all 6 frames. For each query, the genbank ID and the alignment coordinates for the top scoring tblastx hit as well as the top scoring tblastx hit derived from a clinical pathogenic isolate was obtained. A hit was considered to be harbored by a clinical pathogenic isolate if the genbank publication reference specifically denoted a pathogenic clinical isolate as the origin of the sequence. For each of these top scoring hits, any annotated feature which overlapped the alignment coordinates in the specific genbank file was obtained from GenBank. Global sequence alignments and corresponding percentage identities between the query and the obtained sequences were computed using clustalW (*3*) at the nucleotide level as well as the amino acid level in the annotated frame. When multiple annotated features were obtained for a query sequence, only the sequence most similar to the query at the nucleotide level was retained (Tables S3 and S4).

Additionally, the sequence identity between resistance genes isolated from individuals 1 and 2 were computed. For each resistance gene isolated from individual 1, the global sequence identity to the most similar gene (as determined by BLAST) from individual 2 was computed using clustalW. Genes were categorized as either unique to an individual or shared between the two, based on a greater than 90 % global nucleotide sequence identity threshold for similarity (Table S7). Interestingly, more than 65 % of the genes derived from cultured aerobic gut isolates are shared between the two individuals, whereas less than 10 % of the metagenomically derived genes are shared between the two individuals. All of the genes (metagenomic and cultured gut aerobes) shared between individuals have close homologs in GenBank. While 84 % of the genes unique to an individual from cultured aerobic isolates also have close homologs in GenBank, only 11 % of the metagenomically derived genes unique to an individual have close homologs in GenBank. These data argue that application of functional metagenomic selections to the microbiomes of additional individuals would discover a significant number of novel antibiotic resistance genes.

Library sequence coverage and estimation of total number of resistance conferring inserts

Of the 210 inserts, 95 were derived from metagenomic libraries of a total size of 9.3 Gb and 115 were derived from libraries made from cultured aerobes with a total size of 7.5 Gb (Table S5). For the metagenomically derived inserts 58% of these were only sequenced once; for the libraries derived from cultured aerobes this number was 67% (Figure S4 C and D). The fact that more than half of the inserts identified have only been sequenced once suggests that many more resistance genes can be identified in these libraries if more clones are sequenced. In an attempt to estimate the total number of unique resistance conferring inserts in our libraries, we performed a least square fit of a Poisson distribution to our sequence coverage distributions (Figure S4). While the criteria for Poisson statistics (mean equals variance) are poorly satisfied for our data set, we obtained least square fits to our sequence coverage distributions for our metagenomic and cultured gut aerobe libraries with the mean equal to 0.62 and 0.175, respectively (Figure S4 A and B). Based on these mean values and Poisson statistics we would expect to find another 440 resistance conferring inserts by exhaustively sequencing our libraries derived from cultured aerobes and another 89 resistance conferring inserts from our metagenomic libraries.

Naming of novel beta-lactamase genes

For genes that had significant homology (>85%) at the amino acid level to previously known beta-lactamase proteins, names were assigned according to the standard beta-lactamase conventions (*4*) with the family name (*CblA*, *CfxA* and *TEM*) followed by a sequentially increasing number (Table 1 and Figure 4). Numbers were chosen to increment from the latest named enzyme deposited to Genbank from each family (*4*). Due to the specific nomenclature in the literature for the *AmpC* family of enzymes (*5, 6*), novel genes were named with the family followed by the two letter code HG (**H**uman **G**ut) denoting origin of enzyme followed by a number (Table 1 and Figure 4).

Beta-lactamase genes that encoded proteins less than 65 % identical to known gene products in GenBank were considered to be novel beta-lactamase enzyme families. Each family was named according to the guidelines outlined in (*5, 7*), with a two letter code corresponding to the source (HG – Human Gut and HO – Human Oral) followed by a third unique letter distinguishing each family (Table 1 and Figure 4).

Phylogenetic analysis of antibiotic resistance genes

We performed phylogenetic analyses for all genes in our study, categorized by the specific antibiotic class to which they confer resistance*.* The amino acid sequences for unique resistance genes derived from functional selections of genomic and metagenomic libraries from gut and oral microbiomes from healthy humans were aligned in ClustalX (*8*) using the GONNET protein weight matrix. Unrooted phylogenetic trees were generated from ClustalX alignments, using the neighbor-joining algorithm of Saitou and Nei (*9*) based on the principle of minimum-evolution, along with bootstrap analysis (1000 replicates) in ClustalX. For the beta-lactamase resistance genes shown in Figure 4, the tree was plotted using DRAWTREE in the PHYLIP package, and except for the nodes indicated all bootstrap values = 1000. For the remaining resistance gene classes (Figures S5, S6 and S7), trees were plotted using NJtree and all bootstrap values are displayed. Tree branch lengths are proportional to relative sequence identity, and the scale bar is in fixed amino acid substitutions per sequence position.

Limitations of PCR and metagenomic functional selections in sampling resistance reservoirs

Antibiotic resistance genes in microbial environments can be interrogated using PCR based approaches (*10-12*). However, the success of these methods is limited to detection of sequences similar to known genes and requires separate functional confirmation of amplified genes. In comparison, as we have shown, metagenomic functional selections directly associate the selected resistance phenotype with the identified full length gene transferred to the selection host, and are thus ideally suited for interrogation of the microbiomic resistance reservoir that is accessible to lateral gene transfer. However, metagenomic functional selections for characterization of resistance reservoirs have certain limitations as well. These primarily involve the requirement for (1) host susceptibility to the antibiotics assayed, (2) appropriate translation and functional compatibility of heterologous gene products in the host, and (3) the resistance gene to encode a dominant phenotype that overcomes the host mechanism of susceptibility. Methods to overcome these limitations include (1) use of multiple host strains which span the spectrum of antibiotic susceptibility mechanisms assayed, (2) engineering of new host strains with an expanded repertoire of translational machinery including rare codon transfer RNAs, and (3) the use of conditional host strain mutants which suppress specific dominant antibiotic susceptibility loci allowing complementation by a transferred resistance gene that is otherwise recessive. Despite current limitations, metagenomic functional selections are useful for the identification of numerous known as well as novel resistance genes, since their success is independent of prior knowledge of the sequence of previously identified genes (*13*).

Supporting Figures

Figure S1: Phylogenetic distribution of gut microbiome day 1 isolates. The 16S ribosomal DNA (rDNA) of cultured isolates was amplified using universal bacterial 16S primers, and sequenced for phylogenetic profiling. Phylogeny was determined using SeqMatch from the Ribosomal Database Project II server(*1*)

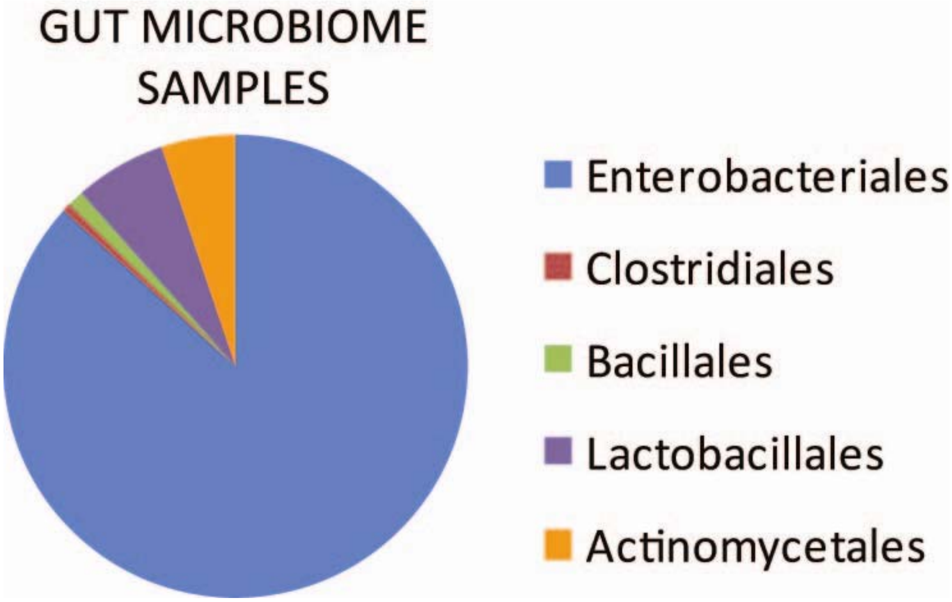


Figure S2: Heat maps of antibiotic resistance of bacterial isolates from gut microbiome 1 at day 1, 140 and 141. Growth measurements of microbiome isolates after 24 hrs at 37 °C in rich media containing antibiotics at concentrations preventing growth of wild type *E.coli* (Table S1) are displayed as linear color-scaled squares, where white denotes no growth, and color intensity is proportional to growth in the presence of antibiotic, scaled to growth in the absence of antibiotic per individual isolate.

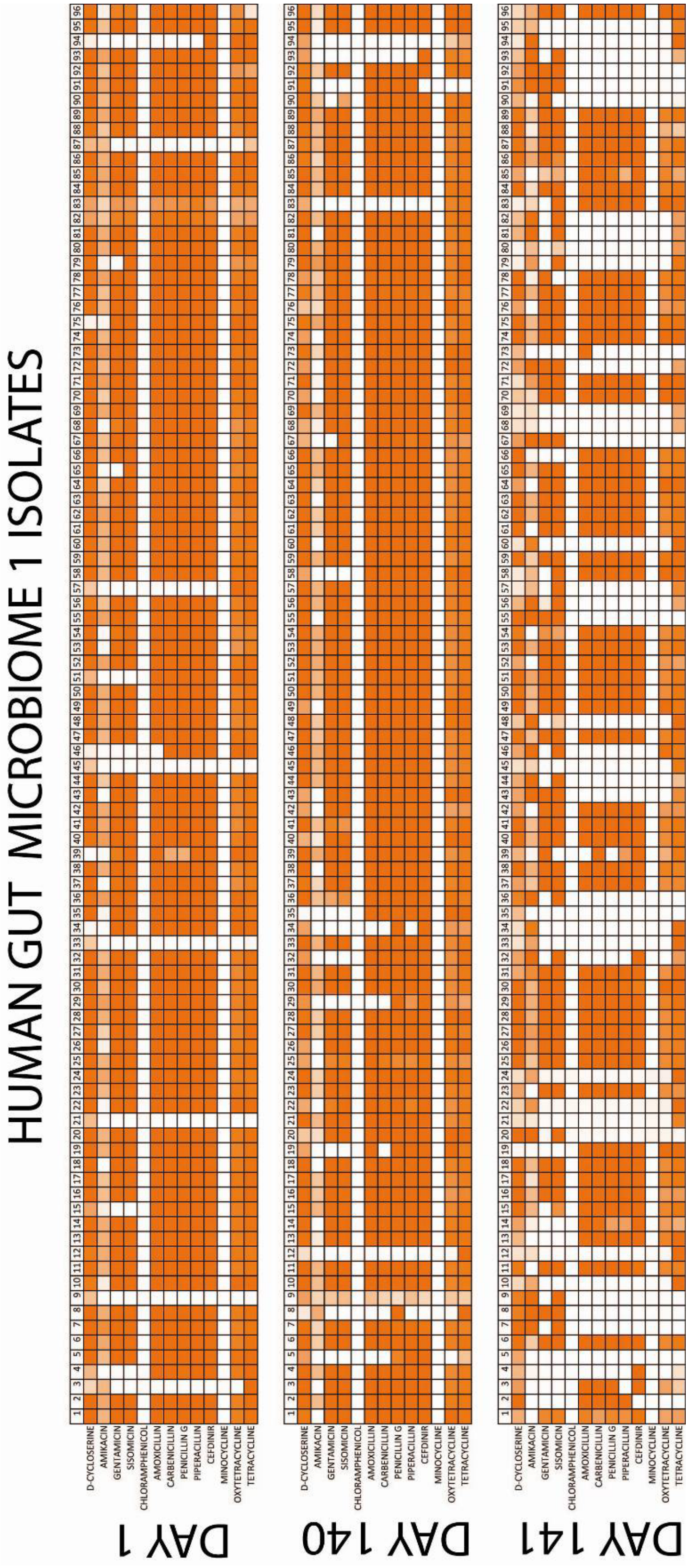

HUMAN GUT MICROBIOME 1 ISOLATES
DAY 1
DAY 140
DAY 141
D-CYCLOSERINE
AMIKACIN
GENTAMICIN
SISOMICIN
CHLORAMPHENICOL
AMOXICILLIN
CARBENICILLIN
PENICILLIN G
PIPERACILLIN
CEFDINIR
MINOCYCLINE
OXYTETRACYCLINE
TETRACYCLINE
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96

Figure S3: Heat maps of antibiotic resistance of bacterial isolates from gut microbiome 2 at day 1, 140 and 141. Growth measurements of microbiome isolates after 24 hrs at 37 °C in rich media containing antibiotics at concentrations preventing growth of wild type *E.coli* (Table S1) are displayed as linear color-scaled squares, where white denotes no growth, and color intensity is proportional to growth in the presence of antibiotic, scaled to growth in the absence of antibiotic per individual isolate.

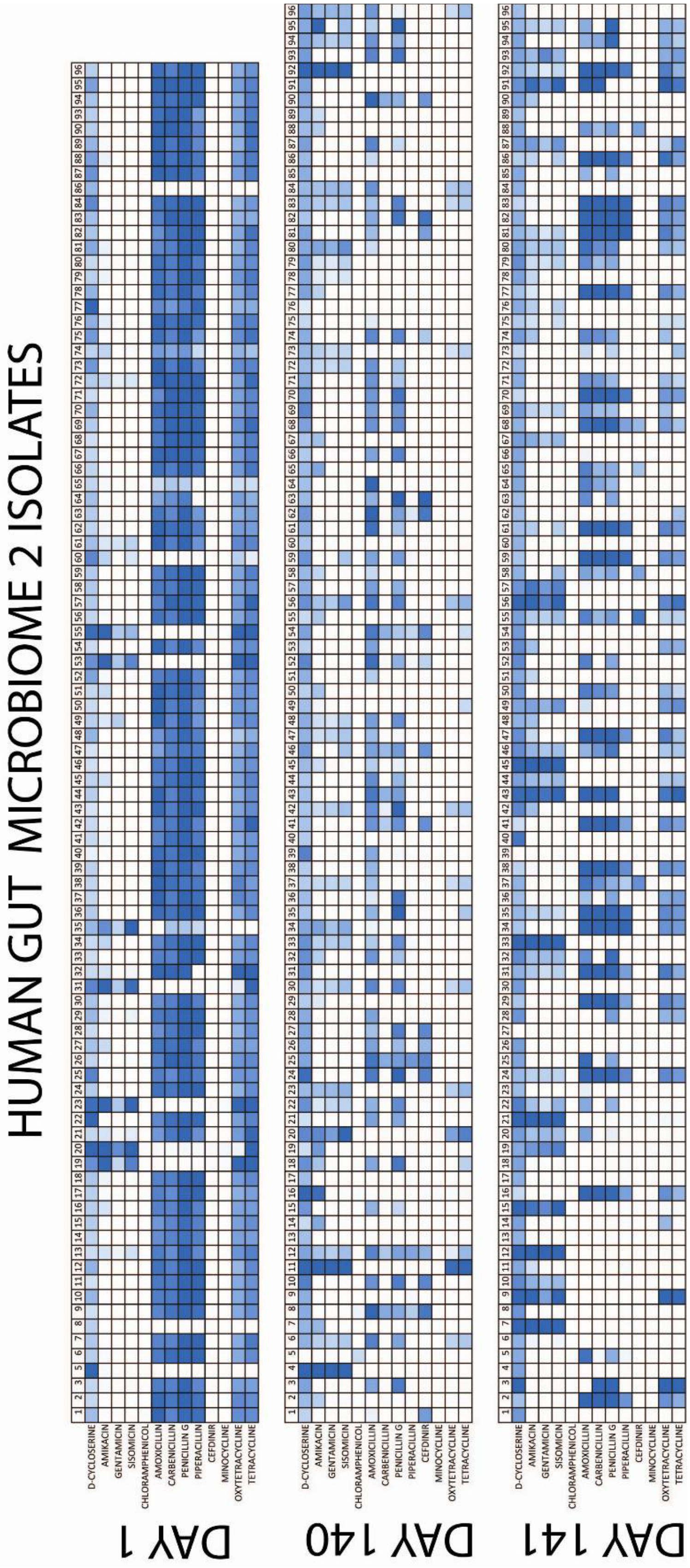

HUMAN GUT MICROBIOME 2 ISOLATES
DAY 1
DAY 140
DAY 141
D-CYCLOSERINE
AMIKACIN
GENTAMICIN
SISOMICIN
CHLORAMPHENICOL
AMOXICILLIN
CARBENICILLIN
PENICILLIN G
PIPERACILLIN
CEFDINIR
MINOCYCLINE
OXYTETRACYCLINE
TETRACYCLINE
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96

Figure S4: Sequencing sampling depth of resistance conferring inserts. For each sampling depth, the number of unique resistance conferring inserts is displayed for (A) metagenomic libraries and (B) the libraries derived from cultured aerobic gut isolates. The dotted line shows the best fit of a Poisson distribution to our sampling data, where the mean of the best Poisson fit for the metagenomic libraries is 0.62 and for the libraries derived from the cultured aerobic isolates is 0.17. The number of times that a particular insert has been sequenced is displayed for (C) the metagenomic libraries and (D) the libraries derived from the cultured aerobic isolates.

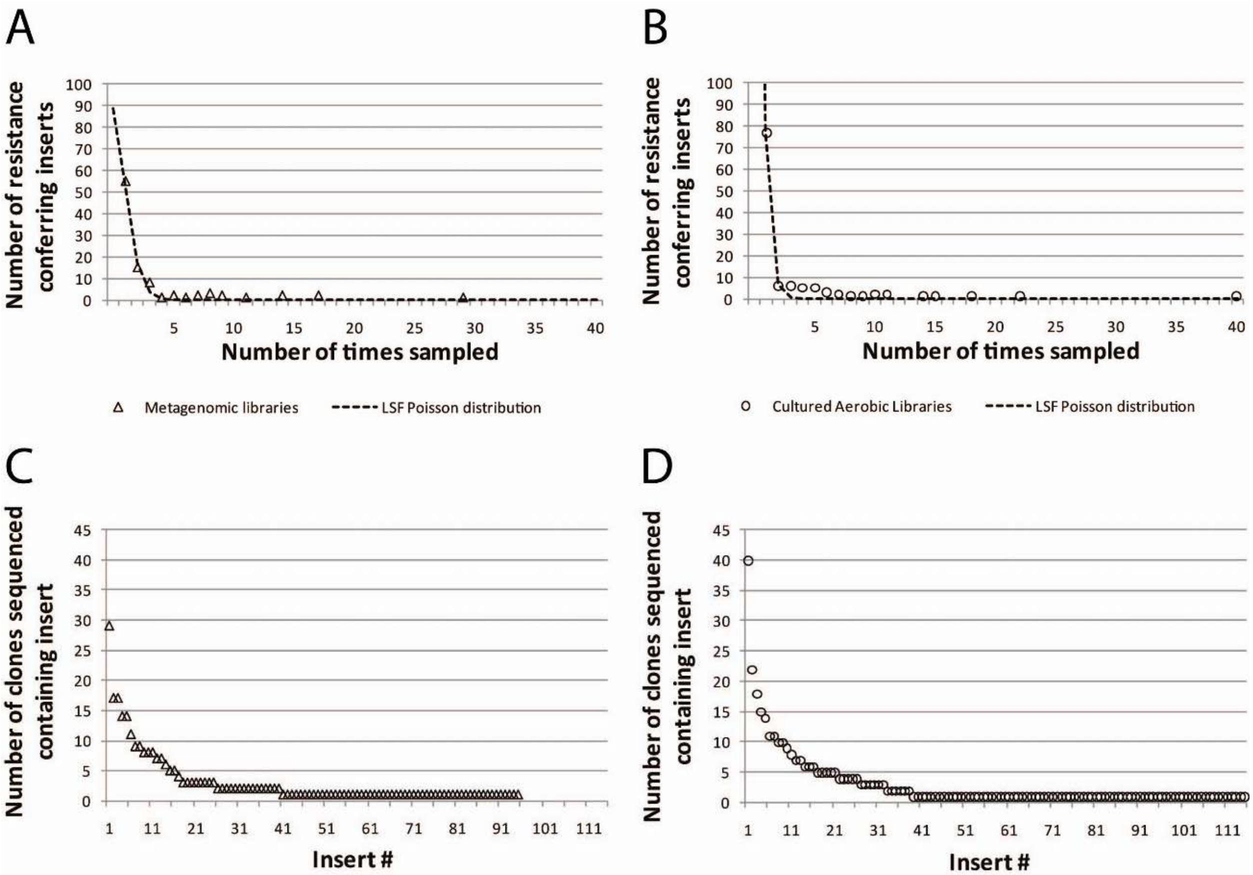


Figure S5: Phylogenetic distribution of full-length gene products encoding aminoglycoside resistance from human microbiomes. The amino acid sequences for all aminoglycoside resistance genes encoded within unique inserts derived from functional selections of genomic (blue labels) and metagenomic (red labels) libraries from gut and oral microbiomes from two healthy humans were aligned in ClustalX (*8*) along with the most closely related genes in GenBank (black labels) (Tables S3 and S4) using the GONNET protein weight matrix. Unrooted phylogenetic trees were generated from ClustalX alignments, using the neighbor-joining algorithm of Saitou and Nei (*9*) based on the principle of minimum-evolution, along with bootstrap analysis (1000 replicates) in ClustalX, and plotted using NJtree. Bootstrap values are indicated on corresponding branches. Tree branch lengths are proportional to relative sequence identity, and the scale bar is in fixed amino acid substitutions per sequence position.

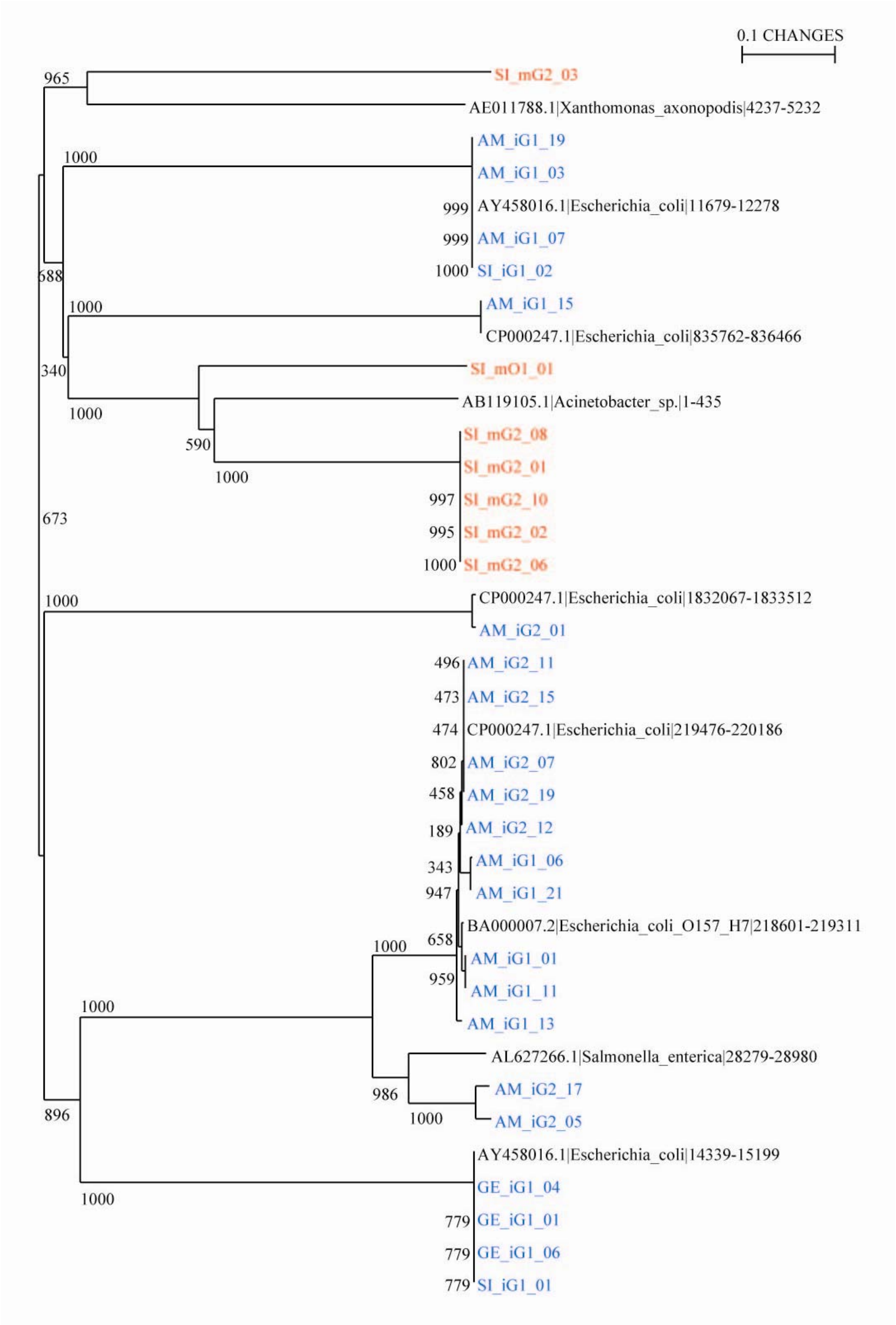
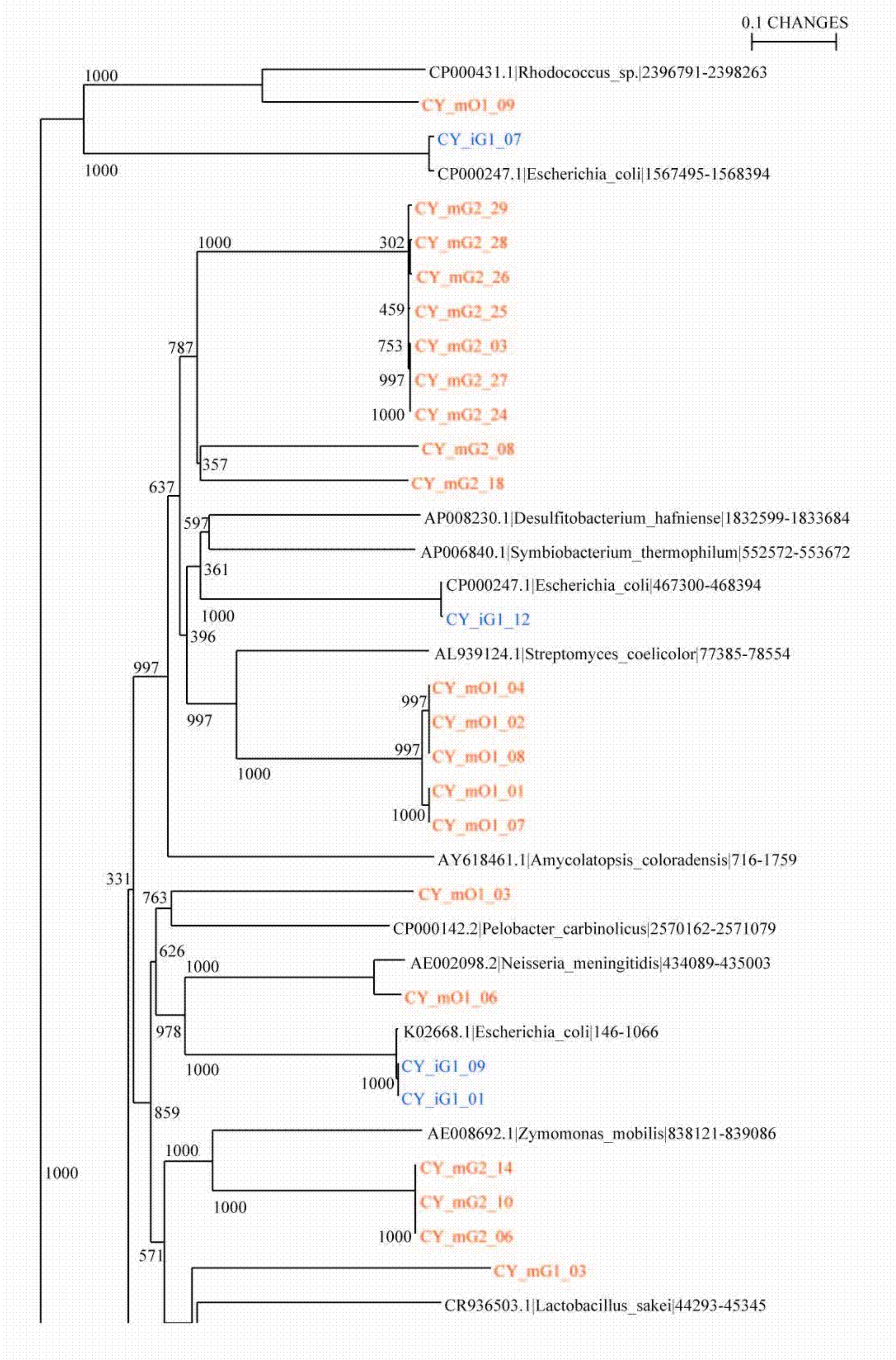

0.1 CHANGES
965
SI_mG2_03
AE011788.1|Xanthomonas_axonopodis|4237-5232
1000
AM_iG1_19
AM_iG1_03
999 AY458016.1|Escherichia_coli|11679-12278
999 AM_iG1_07
1000 SI_iG1_02
688
1000
AM_iG1_15
CP000247.1|Escherichia_coli|835762-836466
340
SI_mO1_01
AB119105.1|Acinetobacter_sp.|1-435
1000
590
SI_mG2_08
SI_mG2_01
1000
997 SI_mG2_10
995 SI_mG2_02
1000 SI_mG2_06
673
1000
CP000247.1|Escherichia_coli|1832067-1833512
AM_iG2_01
496 AM_iG2_11
473 AM_iG2_15
474 CP000247.1|Escherichia_coli|219476-220186
802 AM_iG2_07
458 AM_iG2_19
189 AM_iG2_12
343 AM_iG1_06
947 AM_iG1_21
BA000007.2|Escherichia_coli_O157_H7|218601-219311
1000
658
AM_iG1_01
959
AM_iG1_11
AM_iG1_13
1000
AL627266.1|Salmonella_enterica|28279-28980
986
AM_iG2_17
1000
AM_iG2_05
896
AY458016.1|Escherichia_coli|14339-15199
GE_iG1_04
1000
779 GE_iG1_01
779 GE_iG1_06
779 SI_iG1_01

Figure S6: Phylogenetic distribution of full-length gene products encoding D-cycloserine resistance from human microbiomes. The amino acid sequences for all D-cycloserine resistance genes encoded within unique inserts derived from functional selections of genomic (blue labels) and metagenomic (red labels) libraries from gut and oral microbiomes from two healthy humans were aligned in ClustalX (*8*) along with the most closely related genes in GenBank (black labels) (Tables S3 and S4) using the GONNET protein weight matrix. Unrooted phylogenetic trees were generated from ClustalX alignments, using the neighbor-joining algorithm of Saitou and Nei (*9*) based on the principle of minimum-evolution, along with bootstrap analysis (1000 replicates) in ClustalX, and plotted using NJtree. Bootstrap values are indicated on corresponding branches. Tree branch lengths are proportional to relative sequence identity, and the scale bar is in fixed amino acid substitutions per sequence position.


0.1 CHANGES
1000
CP000431.1|Rhodococcus_sp.|2396791-2398263
CY_mO1_09
CY_iG1_07
CP000247.1|Escherichia_coli|1567495-1568394
1000
CY_mG2_29
1000
302
CY_mG2_28
CY_mG2_26
459
CY_mG2_25
787
753
CY_mG2_03
997
CY_mG2_27
1000
CY_mG2_24
CY_mG2_08
357
637
CY_mG2_18
597
AP008230.1|Desulfitobacterium_hafniense|1832599-1833684
AP006840.1|Symbiobacterium_thermophilum|552572-553672
361
CP000247.1|Escherichia_coli|467300-468394
1000
CY_iG1_12
396
AL939124.1|Streptomyces_coelicolor|77385-78554
997
CY_mO1_04
997
997
CY_mO1_02
997
CY_mO1_08
1000
CY_mO1_01
1000
CY_mO1_07
AY618461.1|Amycolatopsis_coloradensis|716-1759
331
763
CY_mO1_03
CP000142.2|Pelobacter_carbinolicus|2570162-2571079
626
1000
AE002098.2|Neisseria_meningitidis|434089-435003
CY_mO1_06
978
K02668.1|Escherichia_coli|146-1066
1000
CY_iG1_09
1000
CY_iG1_01
859
AE008692.1|Zymomonas_mobilis|838121-839086
1000
CY_mG2_14
1000
1000
CY_mG2_10
1000
CY_mG2_06
571
CY_mG1_03
CR936503.1|Lactobacillus_sakei|44293-45345

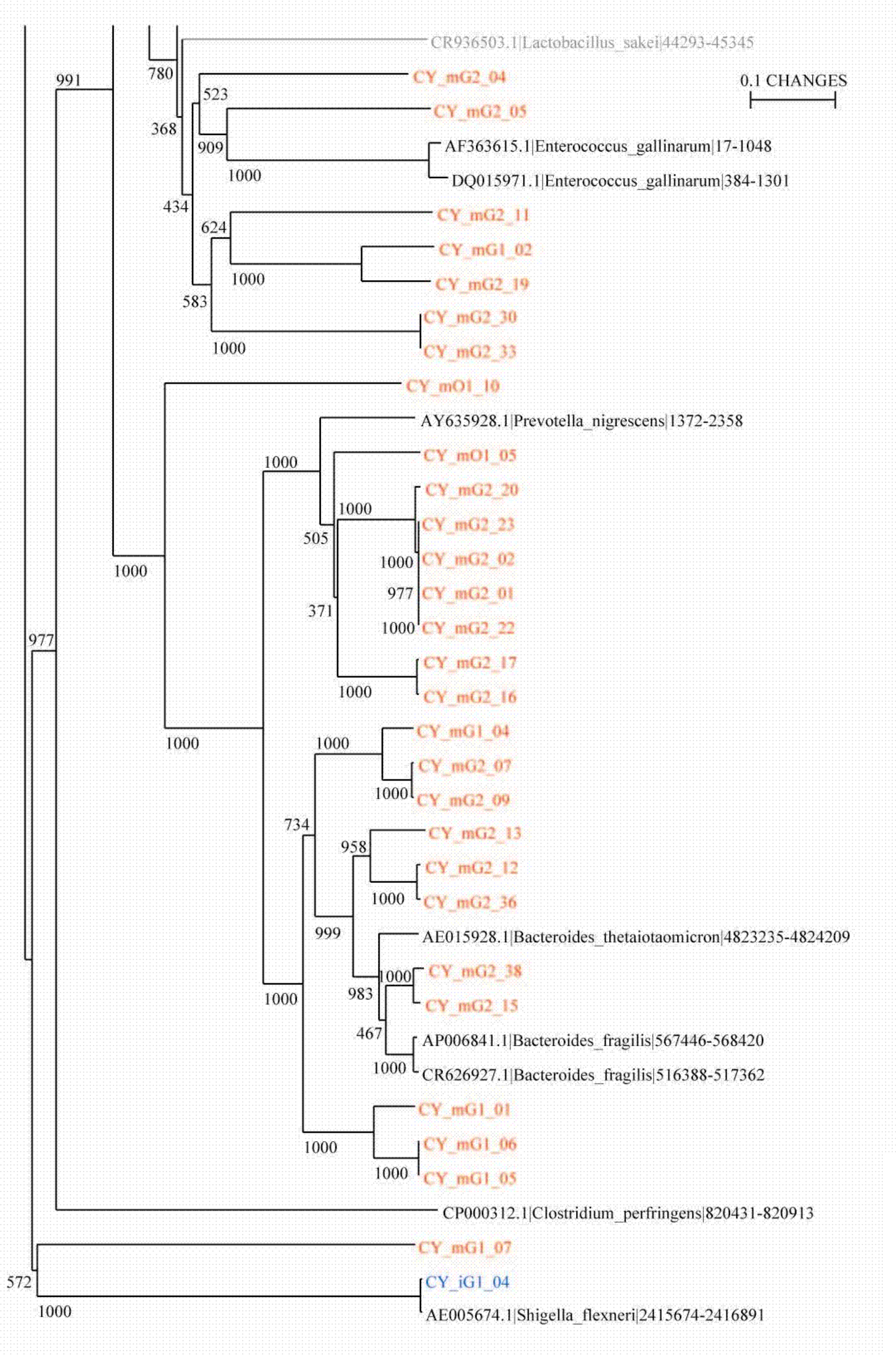
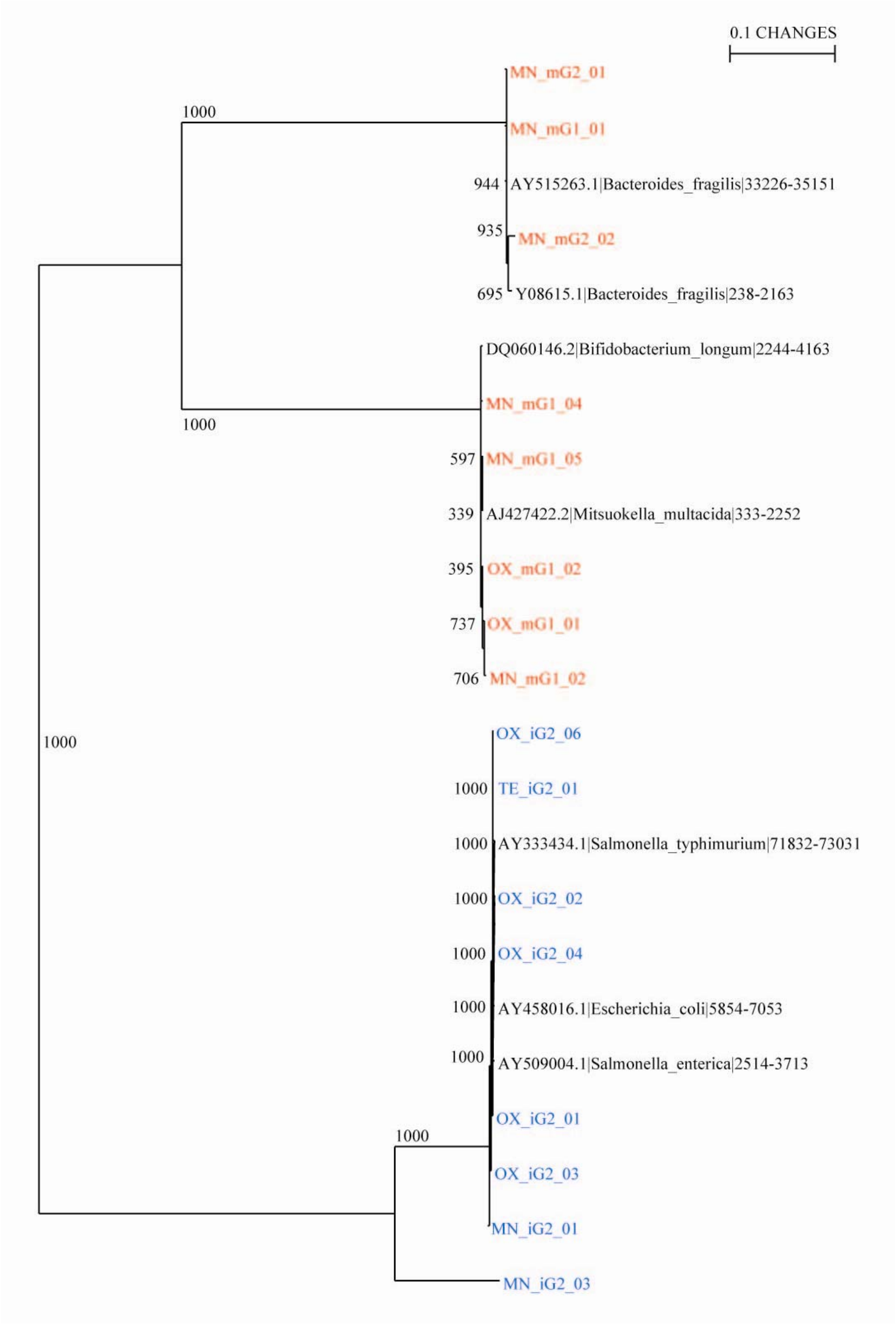

CR936503.1|Lactobacillus_sakei|44293-45345
CY_mG2_04
0.1 CHANGES
CY_mG2_05
AF363615.1|Enterococcus_gallinarum|17-1048
DQ015971.1|Enterococcus_gallinarum|384-1301
CY_mG2_11
CY_mG1_02
CY_mG2_19
CY_mG2_30
CY_mG2_33
CY_mO1_10
AY635928.1|Prevotella_nigrescens|1372-2358
CY_mO1_05
CY_mG2_20
CY_mG2_23
CY_mG2_02
CY_mG2_01
CY_mG2_22
CY_mG2_17
CY_mG2_16
CY_mG1_04
CY_mG2_07
CY_mG2_09
CY_mG2_13
CY_mG2_12
CY_mG2_36
AE015928.1|Bacteroides_thetaiotaomicron|4823235-4824209
CY_mG2_38
CY_mG2_15
AP006841.1|Bacteroides_fragilis|567446-568420
CR626927.1|Bacteroides_fragilis|516388-517362
CY_mG1_01
CY_mG1_06
CY_mG1_05
CP000312.1|Clostridium_perfringens|820431-820913
CY_mG1_07
CY_iG1_04
AE005674.1|Shigella_flexneri|2415674-2416891
991
780
523
368
909
1000
434
624
1000
583
1000
1000
1000
505
1000
1000
977
371
1000
1000
1000
1000
1000
734
958
1000
999
1000
983
467
1000
1000
1000
977
572
1000

Figure S7: Phylogenetic distribution of full-length gene products encoding resistance to tetracycline antibiotics from human microbiomes. The amino acid sequences for all tetracycline resistance genes encoded within unique inserts derived from functional selections of genomic (blue labels) and metagenomic (red labels) libraries from gut and oral microbiomes from two healthy humans were aligned in ClustalX (*8*) along with the most closely related genes in GenBank (black labels) (Tables S3 and S4) using the GONNET protein weight matrix. Unrooted phylogenetic trees were generated from ClustalX alignments, using the neighbor-joining algorithm of Saitou and Nei (*9*) based on the principle of minimum-evolution, along with bootstrap analysis (1000 replicates) in ClustalX, and plotted using NJtree. Bootstrap values are indicated on corresponding branches. Tree branch lengths are proportional to relative sequence identity, and the scale bar is in fixed amino acid substitutions per sequence position.


0.1 CHANGES
1000
MN_mG2_01
MN_mG1_01
944 AY515263.1|Bacteroides_fragilis|33226-35151
935 MN_mG2_02
695 Y08615.1|Bacteroides_fragilis|238-2163
DQ060146.2|Bifidobacterium_longum|2244-4163
MN_mG1_04
1000
597 MN_mG1_05
339 AJ427422.2|Mitsuokella_multacida|333-2252
395 OX_mG1_02
737 OX_mG1_01
706 MN_mG1_02
1000
OX_iG2_06
1000 TE_iG2_01
1000 AY333434.1|Salmonella_typhimurium|71832-73031
1000 OX_iG2_02
1000 OX_iG2_04
1000 AY458016.1|Escherichia_coli|5854-7053
1000 AY509004.1|Salmonella_enterica|2514-3713
OX_iG2_01
1000
OX_iG2_03
MN_iG2_01
MN_iG2_03

Supporting Tables

Table S1: Antibiotics used for functional selections of antibiotic resistance machinery from human microbiomes. Concentrations are chosen that prevent growth of *E. coli* TOP10 (Invitrogen)

Antibiotic	**Code**	**Concentration (ug/mL)**	**Class**	**Type**
D-Cycloserine	CY	100	Amino acid derivative	Natural
Amikacin	AM	50	Aminoglycoside	Semi-synthetic
Gentamicin	GE	50	Aminoglycoside	Natural
Sisomicin	SI	20	Aminoglycoside	Natural
Chloramphenicol	CH	20	Amphenicol	Natural
Amoxicillin	AX	20	Beta-Lactam	Semi-synthetic
Carbenicillin	CA	50	Beta-Lactam	Semi-synthetic
Penicillin G	PE	50	Beta-Lactam	Natural
Piperacillin	PI	50	Beta-Lactam	Semi-synthetic
Cefdinir	CF	20	Cephalosporin	Semi-synthetic
Minocycline	MN	20	Tetracyclines	Semi-synthetic
Oxytetracycline	OX	20	Tetracyclines	Natural
Tetracycline	TE	20	Tetracyclines	Natural

Table S2: Selected genbank IDs of entries with identified *TEM1* variant and the strain from which it was derived.

Genbank ID	**Pathogenic strain**
EU780013.1	*Klebsiella pneumoniae* strain 12836
EU219534.1	*Salmonella enterica* subsp. enterica serovar
AB070224.1	*Serratia marcescens*
AM849806.1	*Haemophilus parainfluenzae*
EU213261.1	*Escherichia coli*
AM884335.1	*Haemophilus influenzae*
AF126482.1	*Neisseria meningitidis*
X54607.1	*Pseudomonas aeruginosa*

Table S3: Resistance genes identified using metagenomic functional selections from human microbiomes. Gene ID is made up of the 2 letter code for the antibiotics used for the selections (e.g. AX denotes Amoxicillin, see Table S1), an m denoting metagenomic origin of DNA insert, microbiome two letter code (e.g. O1 denotes oral microbiome 1, G2 denotes gut microbiome 2) and) and a numeric identifier. For each gene identified, the most similar gene from any organism as well as the most similar gene harbored by a pathogenic isolate in GenBank was identified using tblastx. Global sequence identities at the nucleotide and amino acid level between these genes were computed using clustalW (*3*).

Gene ID	**Gene length [bp]**	**Genbank ID**	**Gene Annotation**	**Top hit [gbID|title|position]**	**Global DNA % ID**	**Global Protein % ID**	**Top pathogenic hit [gbID|title|position]**	**Global DNA % ID**	**Global Protein % ID**
AX_mG1_01	603	GQ342997	beta-lactamase	L08472.1|Bacteroides_uniformis|848-1738	50.8	35.7	AF148067.1|Escherichia_coli|214-1158	56.9	35.5
AX_mG1_04	891	GQ342998	beta-lactamase	L08472.1|Bacteroides_uniformis|848-1738	99.9	100.0	AF148067.1|Escherichia_coli|214-1158	53.3	40.7
AX_mG2_01	996	GQ342996	CfxA2-like beta-lactamase	AF472622.2|Capnocytophaga_ochracea|53-1018	89.1	87.2	AF472622.2|Capnocytophaga_ochracea|53-1018	89.1	87.2
AX_mG2_03	891	GQ342999	beta-lactamase	L08472.1|Bacteroides_uniformis|848-1738	99.9	99.7	AF148067.1|Escherichia_coli|214-1158	53.4	41.0
AX_mG2_05	906	GQ343000	beta-lactamase	L08472.1|Bacteroides_uniformis|848-1738	61.9	58.5	AF148067.1|Escherichia_coli|214-1158	54.1	42.2
AX_mG2_07	996	GQ343001	CfxA2-like beta-lactamase	AF472622.2|Capnocytophaga_ochracea|53-1018	89.1	87.2	AF472622.2|Capnocytophaga_ochracea|53-1018	89.1	87.2
AX_mG2_09	873	GQ343002	beta-lactamase	L08472.1|Bacteroides_uniformis|848-1738	55.1	43.3	AF148067.1|Escherichia_coli|214-1158	52.2	37.9
AX_mG2_11	891	GQ343003	beta-lactamase	L08472.1|Bacteroides_uniformis|848-1738	53.3	37.1	AF339733.1|Chryseobacterium_gleum|437-1315	47.0	34.3
AX_mO1_01	882	GQ343035	beta-lactamase	AF339733.1|Chryseobacterium_gleum|437-1315	59.6	49.5	AF339733.1|Chryseobacterium_gleum|437-1315	59.6	49.5
CA_mG1_01	894	GQ343037	beta-lactamase	L13472.1|Bacteroides_fragilis|436-1338	57.8	48.1	L13472.1|Bacteroides_fragilis|436-1338	57.8	48.1

Gene ID	Gene length [bp]	Genbank ID	Gene Annotation	Top hit [gbID|title|position]	Global DNA % ID	Global Protein % ID	Top pathogenic hit [gbID|title|position]	Global DNA % ID	Global Protein % ID
CA_mG1_02	903	GQ343038	beta-lactamase	L08472.1|Bacteroides_uniformis|848-1738	65.2	61.4	AF148067.1|Escherichia_coli|214-1158	54.7	43.3
CA_mG1_04	894	GQ343039	beta-lactamase	AE015928.1|Bacteroides_thetaiotaomicron_VPI-5482|5922082-5922963	57.7	51.0	AF504912.1|Prevotella_denticola|1-966	50.8	39.4
CA_mG1_05	894	GQ343041	beta-lactamase	L13472.1|Bacteroides_fragilis|436-1338	57.8	48.1	L13472.1|Bacteroides_fragilis|436-1338	57.8	48.1
CA_mG1_07	891	GQ343040	class A beta lactamase	L08472.1|Bacteroides_uniformis|848-1738	99.9	100.0	X93314.1|Salmonella_typhimurium|278-1204	50.3	37.4
CA_mG1_09	894	GQ343042	class A beta-lactamase	L13472.1|Bacteroides_fragilis|436-1338	57.8	48.1	L13472.1|Bacteroides_fragilis|436-1338	57.8	48.1
CA_mG2_01	822	GQ343043	beta-lactamase	L08472.1|Bacteroides_uniformis|848-1738	99.9	100.0	AF148067.1|Escherichia_coli|214-1158	54.1	43.0
CA_mG2_04	960	GQ343044	beta-lactamase	AE015928.1|Bacteroides_thetaiotaomicron_VPI-5482|5922082-5922963	59.1	52.9	AF148067.1|Escherichia_coli|214-1158	54.5	39.9
CA_mG2_07	888	GQ343045	metallo-beta-lactamase	AY372695.1|Bacteroides_fragilis|1787-2536	53.2	42.6	AE016879.1|Bacillus anthracis str. Ames|3216832-3217602	49.0	30.5
CA_mG2_10	495	GQ343036	CfxA beta-lactamase	AY769933.1|Bacteroides_fragilis|1-966	88.9	87.9	AF472622.2|Capnocytophaga_ochracea|53-1018	88.7	87.3
CA_mG2_12	891	GQ343046	class A beta lactamase	L08472.1|Bacteroides_uniformis|848-1738	53.3	37.1	AF148067.1|Escherichia_coli|214-1158	47.3	30.8
CA_mO1_02	459	GQ343066	tetracycline resistance protein D	BX664015.1|Serratia_marcescens|204138-204554	100.0	100.0	BX664015.1|Serratia_marcescens|204138-204554	100.0	100.0
CF_mG2_01	891	GQ343068	beta-lactamase	L08472.1|Bacteroides_uniformis|848-1738	99.9	100.0	AF148067.1|Escherichia_coli|214-1158	53.4	40.7
CF_mG2_04	891	GQ343067	class A beta lactamase	L08472.1|Bacteroides_uniformis|848-1738	99.9	100.0	AF148067.1|Escherichia_coli|214-1158	53.4	40.7
CH_mG2_05	624	GQ343079	chloramphenicol O-acetyltransferase	AF031037.1|Neisseria_meningitidis|489-1112	82.9	81.2	AF031037.1|Neisseria_meningitidis|489-1112	82.9	81.2
CH_mG2_07	624	GQ343078	chloramphenicol O-acetyltransferase	AF031037.1|Neisseria_meningitidis|489-1112	82.9	81.2	AF031037.1|Neisseria_meningitidis|489-1112	82.9	81.2

CY_mG1_01	981	GQ343082	D-alanyl-alanine synthetase A	AP006841.1|Bacteroides_fragilis|567446-568420	68.7	73.1	AE015924.1|Porphyromonas_gingivalis|779916-780908	58.1	49.4
Gene ID	**Gene length [bp]**	**Gen-bank ID**	**Gene Annotation**	**Top hit [gbID|title|position]**	**Global DNA % ID**	**Global Protein % ID**	**Top pathogenic hit [gbID|title|position]**	**Global DNA % ID**	**Global Protein % ID**
CY_mG1_02	978	GQ343084	D-alanine--D-alanine ligase	AF363615.1|Enterococcus_gallinarum|17-1048	54.7	43.3	AF363615.1|Enterococcus_gallinarum|17-1048	54.7	43.3
CY_mG1_03	627	GQ343085	D-alanine--D-alanine ligase	CR936503.1|Lactobacillus_sakei_subsp.|44293-45345	52.1	31.9	AE004091.2|Pseudomonas_aeruginosa|4700906-4701946	41.1	30.5
CY_mG1_04	966	GQ343086	D-alanyl-alanine synthetase A	AP006841.1|Bacteroides_fragilis|567446-568420	71.4	78.9	AP006841.1|Bacteroides_fragilis|567446-568420	71.4	78.9
CY_mG1_05	978	GQ343087	D-alanyl-alanine synthetase A	CR626927.1|Bacteroides_fragilis_NCTC_9343|516388-517362	68.8	72.2	CR626927.1|Bacteroides_fragilis_NCTC_9343|516388-517362	68.8	72.2
CY_mG1_06	978	GQ343088	D-alanyl-alanine synthetase A	CR626927.1|Bacteroides_fragilis_NCTC_9343|516388-517362	68.8	72.2	CR626927.1|Bacteroides_fragilis_NCTC_9343|516388-517362	68.8	72.2
CY_mG1_07	885	GQ343083	aminotransferase, class I/II	CP000312.1|Clostridium_perfringens|820910-822088	52.5	54.3	CP000312.1|Clostridium_perfringens|820910-822088	52.5	54.3
CY_mG2_01	996	GQ343080	D-alanyl-alanine synthetase A	AY635928.1|Prevotella_nigrescens|1372-2358	69.6	78.4	AY635928.1|Prevotella_nigrescens|1372-2358	69.6	78.4
CY_mG2_02	996	GQ343089	D-alanyl-alanine synthetase A	AY635928.1|Prevotella_nigrescens|1372-2358	69.6	78.4	AY635928.1|Prevotella_nigrescens|1372-2358	69.6	78.4
CY_mG2_03	1065	GQ343092	D-alanine--D-alanine ligase	AP008230.1|Desulfitobacterium_hafniense|1832599-1833684	55.7	43.8	CP000001.1|Bacillus_cereus_E33L_Bacteria;_Firmicutes;_Bacillales;_Bacillaceae;_Bacillus;_Bacillus|235218-236303	53.5	43.2
CY_mG2_04	1032	GQ343099	D-alanine--D-alanine ligase	AF363615.1|Enterococcus_gallinarum|17-1048	52.1	46.0	AF363615.1|Enterococcus_gallinarum|17-1048	52.1	46.0
CY_mG2_05	1047	GQ343100	D-alanine--D-alanine ligase	AF363615.1|Enterococcus_gallinarum|17-1048	53.6	47.8	AF363615.1|Enterococcus_gallinarum|17-1048	53.6	47.8
CY_mG2_06	906	GQ343101	D-alanine--D-alanine ligase	AE008692.1|Zymomonas_mobilis_subsp.|838121-839086	59.3	49.0			
CY_mG2_07	975	GQ343103	D-alanine--D-alanine ligase	AP006841.1|Bacteroides_fragilis|567446-568420	71.7	78.1	AP006841.1|Bacteroides_fragilis|567446-568420	71.7	78.1

Gene ID	Gene length [bp]	Gen-bank ID	Gene Annotation	Top hit [gbID|title|position]	Global DNA % ID	Global Protein % ID	Top pathogenic hit [gbID|title|position]	Global DNA % ID	Global Protein % ID
CY_mG2_08	1101	GQ343104	D-alanine--D-alanine ligase	AP006840.1|Symbiobacterium_thermophilum_IAM_14863_Bacteria;_Firmicutes;_Clostridia;_Clostridiales;_Clostridiales|552572-553672	59.9	45.7	AE015927.1|Clostridium_tetani|275925-277001	44.4	41.9
CY_mG2_09	975	GQ343105	D-alanine--D-alanine ligase	AP006841.1|Bacteroides_fragilis|567446-568420	70.8	78.1	AP006841.1|Bacteroides_fragilis|567446-568420	70.8	78.1
CY_mG2_10	906	GQ343106	D-alanine--D-alanine ligase	AE008692.1|Zymomonas_mobilis_subsp.|838121-839086	59.3	49.0			
CY_mG2_11	1149	GQ343107	D-alanine--D-alanine ligase	AF363615.1|Enterococcus_gallinarum|17-1048	54.3	42.9	AF363615.1|Enterococcus_gallinarum|17-1048	54.3	42.9
CY_mG2_12	975	GQ343108	D-alanine--D-alanine ligase	AE015928.1|Bacteroides_thetaiotaomicron_VPI-5482|4823235-4824209	76.2	84.0	AY635928.1|Prevotella_nigrescens|1372-2358	62.6	63.6
CY_mG2_13	981	GQ343109	D-alanine--D-alanine ligase	AE015928.1|Bacteroides_thetaiotaomicron_VPI-5482|4823235-4824209	77.2	86.1	AY635928.1|Prevotella_nigrescens|1372-2358	61.7	62.0
CY_mG2_14	906	GQ343110	D-alanine--D-alanine ligase	AE008692.1|Zymomonas_mobilis_subsp.|838121-839086	59.3	49.0			
CY_mG2_15	1026	GQ343113	D-alanine--D-alanine ligase	AE015928.1|Bacteroides_thetaiotaomicron_VPI-5482|4823235-4824209	80.4	92.3	AY635928.1|Prevotella_nigrescens|1372-2358	63.3	63.6
CY_mG2_16	999	GQ343114	D-alanine--D-alanine ligase	AY635928.1|Prevotella_nigrescens|1372-2358	67.1	79.9	AY635928.1|Prevotella_nigrescens|1372-2358	67.1	79.9
CY_mG2_17	999	GQ343116	D-alanine--D-alanine ligase	AY635928.1|Prevotella_nigrescens|1372-2358	67.2	79.9	AY635928.1|Prevotella_nigrescens|1372-2358	67.2	79.9
CY_mG2_18	1071	GQ343117	D-alanine--D-alanine ligase	AY618461.1|Amycolatopsis_coloradensis|716-1759	59.5	44.5	AE015927.1|Clostridium_tetani|275925-277001	51.4	46.7
CY_mG2_19	1059	GQ343118	D-alanine--D-alanine ligase	AF363615.1|Enterococcus_gallinarum|17-1048	53.0	43.7	AF363615.1|Enterococcus_gallinarum|17-1048	53.0	43.7
CY_mG2_20	996	GQ343081	D-alanyl-alanine synthetase A	AY635928.1|Prevotella_nigrescens|1372-2358	69.3	77.4	AY635928.1|Prevotella_nigrescens|1372-2358	69.3	77.4
CY_mG2_22	996	GQ343090	D-alanyl-alanine synthetase A	AY635928.1|Prevotella_nigrescens|1372-2358	69.6	78.4	AY635928.1|Prevotella_nigrescens|1372-2358	69.6	78.4

Gene ID	Gene length [bp]	Gen-bank ID	Gene Annotation	Top hit [gbID|title|position]	Global DNA % ID	Global Protein % ID	Top pathogenic hit [gbID|title|position]	Global DNA % ID	Global Protein % ID
CY_mG2_23	980	GQ343091	D-alanyl-alanine synthetase A	AY635928.1|Prevotella_nigrescens|1372-2358	69.6	78.4	AY635928.1|Prevotella_nigrescens|1372-2358	69.6	78.4
CY_mG2_24	1065	GQ343093	D-alanine--D-alanine ligase	AP008230.1|Desulfitobacterium_hafniense|1832599-1833684	55.7	43.8	AE016879.1|Bacillus anthracis str. Ames|232804-233079	55.8	41.0
CY_mG2_25	1065	GQ343094	D-alanine--D-alanine ligase	AP008230.1|Desulfitobacterium_hafniense|1832599-1833684	55.8	43.8	AE016879.1|Bacillus anthracis str. Ames|232804-233079	56.2	41.0
CY_mG2_26	1065	GQ343095	D-alanine--D-alanine ligase	AP008230.1|Desulfitobacterium_hafniense|1832599-1833684	55.6	43.8	AE016879.1|Bacillus anthracis str. Ames|232804-233079	57.0	42.2
CY_mG2_27	1065	GQ343096	D-alanine--D-alanine ligase	AP008230.1|Desulfitobacterium_hafniense|1832599-1833684	55.7	43.8	AE016879.1|Bacillus anthracis str. Ames|232804-233079	55.8	41.0
CY_mG2_28	1065	GQ343097	D-alanine--D-alanine ligase	AP008230.1|Desulfitobacterium_hafniense|1832599-1833684	55.6	43.8	AE016879.1|Bacillus anthracis str. Ames|232804-233079	56.2	42.2
CY_mG2_29	1065	GQ343098	D-alanine--D-alanine ligase	AP008230.1|Desulfitobacterium_hafniense|1832599-1833684	55.5	43.8	AE016879.1|Bacillus anthracis str. Ames|232804-233079	56.2	42.2
CY_mG2_30	1104	GQ343102	D-alanine--D-alanine ligase	DQ015971.1|Enterococcus_gallinarum|270-1301	58.6	41.3	AE016879.1|Bacillus anthracis str. Ames|232804-233079	58.6	41.3
CY_mG2_33	1104	GQ343111	D-alanine--D-alanine ligase	DQ015971.1|Enterococcus_gallinarum|270-1301	58.6	41.3	DQ015971.1|Enterococcus_gallinarum|270-1301	58.6	41.3
CY_mG2_36	975	GQ343112	D-alanyl-alanine synthetase A	AE015928.1|Bacteroides_thetaiotaomicron_VPI-5482|4823235-4824209	76.4	84.6	AY635928.1|Prevotella_nigrescens|1372-2358	62.8	64.2
CY_mG2_38	1032	GQ343115	D-alanyl-alanine synthetase A	AE015928.1|Bacteroides_thetaiotaomicron_VPI-5482|4823235-4824209	80.0	91.4	AY635928.1|Prevotella_nigrescens|1372-2358	62.9	62.7
CY_mO1_01	1119	GQ343124	D-alanine--D-alanine ligase	AL939124.1|Streptomyces_coelicolor_A3(2)|77385-78554	64.5	53.4	AE017283.1|Propionibacterium_acnes|1483346-1484443	59.0	50.3
CY_mO1_02	1119	GQ343126	D-alanine--D-alanine ligase	AL939124.1|Streptomyces_coelicolor_A3(2)|77385-78554	64.4	53.1	AE017283.1|Propionibacterium_acnes|1483346-1484443	58.9	50.5

Gene ID	Gene length [bp]	Gen-bank ID	Gene Annotation	Top hit [gbID|title|position]	Global DNA % ID	Global Protein % ID	Top pathogenic hit [gbID|title|position]	Global DNA % ID	Global Protein % ID
CY_mO1_03	933	GQ343127	D-alanine--D-alanine ligase	CP000142.2|Pelobacter_carbinolicus_DSM_2380|2570162-2571079	53.7	44.7	AE006049.1|Pasteurella_multocida_subsp.|2602-3531	51.4	39.0
CY_mO1_04	1119	GQ343128	D-alanine--D-alanine ligase	AL939124.1|Streptomyces_coelicolor_A3(2)|77385-78554	64.4	53.1	AE017283.1|Propionibacterium_acnes|1483346-1484443	58.9	50.5
CY_mO1_05	1005	GQ343129	D-alanine--D-alanine ligase	AY635928.1|Prevotella_nigrescens|1372-2358	68.2	74.7	AY635928.1|Prevotella_nigrescens|1372-2358	68.2	74.7
CY_mO1_06	915	GQ343133	D-alanine--D-alanine ligase	AE002098.2|Neisseria_meningitidis|434089-435003	84.3	93.1	AE002098.2|Neisseria_meningitidis|434089-435003	84.3	93.1
CY_mO1_07	1119	GQ343125	D-alanine--D-alanine ligase	AL939124.1|Streptomyces_coelicolor_A3(2)|77385-78554	64.5	53.4	AB186053.1|Enterococcus_faecalis|1-1071	48.6	41.3
CY_mO1_08	1119	GQ343130	D-alanine--D-alanine ligase	AL939124.1|Streptomyces_coelicolor_A3(2)|77385-78554	64.4	53.1	M20793.1|Salmonella_typhimurium|100-1194	52.7	39.7
CY_mO1_09	1347	GQ343131	D-alanine/D-serine/glycine permease	CP000431.1|Rhodococcus_sp.|2396791-2398263	67.6	61.2	AE005174.2|Escherichia_coli_O157:H7_EDL933|5315324-5316736	60.5	58.2
CY_mO1_10	978	GQ343132	D-alanyl-alanine synthetase A	AP006841.1|Bacteroides_fragilis|567446-568420	54.4	41.9	AP006841.1|Bacteroides_fragilis|567446-568420	54.4	41.9
MN_mG1_01	1926	GQ343137	TetQ	AY515263.1|Bacteroides_fragilis|33226-35151	100.0	100.0	AY515263.1|Bacteroides_fragilis|33226-35151	100.0	100.0
MN_mG1_02	888	GQ343139	TetW	AJ427422.2|Mitsuokella_multacida|333-2252	100.0	100.0	AY466395.1 |Streptococcus pneumoniae|4-1923	65.8	75.6
MN_mG1_04	1920	GQ343140	TetW	DQ060146.2|Bifidobacterium_longum|2244-4163	99.9	99.7	AY466395.1 |Streptococcus pneumoniae|4-1923	63.7	68.1
MN_mG1_05	1920	GQ343138	Tetracycline resistance protein tetW	DQ060146.2|Bifidobacterium_longum|2244-4163	99.9	99.8	Y07780.1|Streptococcus_pneumoniae|125-2044	65.2	67.9
MN_mG2_01	1926	GQ343141	TetQ	Y08615.1|Bacteroides_fragilis|238-2163	100.0	100.0	L42544.1|Streptococcus_pyogenes|478-2433	56.1	48.3
MN_mG2_02	1656	GQ343142	TetA(Q)	Y08615.1|Bacteroides_fragilis|238-2163	99.8	99.6	L42544.1|Streptococcus_pyogenes|478-2433	54.6	44.5

Gene ID	Gene length [bp]	Gen-bank ID	Gene Annotation	Top hit [gbID|title|position]	Global DNA % ID	Global Protein % ID	Top pathogenic hit [gbID|title|position]	Global DNA % ID	Global Protein % ID
OX_mG1_01	1920	GQ343145	Tetracycline resistance protein tetW	DQ060146.2|Bifidobacterium_longum|2244-4163	99.6	99.4	Y07780.1|Streptococcus_pneumoniae|125-2044	65.0	67.4
OX_mG1_02	1920	GQ343146	Tetracycline resistance protein tetW	DQ060146.2|Bifidobacterium_longum|2244-4163	99.6	99.5	Y07780.1|Streptococcus_pneumoniae|125-2044	65.0	67.6
PE_mG1_01	921	GQ343153	class A beta lactamase	L08472.1|Bacteroides_uniformis|848-1738	51.7	38.8	AF148067.1|Escherichia_coli|214-1158	55.0	35.5
PE_mG2_02	891	GQ343154	class A beta-lactamase	L08472.1|Bacteroides_uniformis|848-1738	99.2	99.0	AF148067.1|Escherichia_coli|214-1158	53.3	40.7
PI_mG1_01	909	GQ343170	hypothetical protein - virulence factor	BX571856.1|Staphylococcus_aureus_subsp.|1618968-1621415	46.8	34.5	BX571856.1|Staphylococcus_aureus_subsp.|1618968-1621415	46.8	34.5
PI_mG2_01	876	GQ343169	extended-spectrum beta-lactamase CTX-M-15	DQ485310.1|Escherichia_coli|38-913	100.0	100.0	DQ485310.1|Escherichia_coli|38-913	100.0	100.0
SI_mG2_01	432	GQ343182	6'-N-acetyltransferase	AB119105.1|Acinetobacter_genomosp.|1-435	53.2	47.6	AB119105.1|Acinetobacter_genomosp.|1-435	53.2	47.6
SI_mG2_02	432	GQ343178	6'-N-acetyltransferase	AB119105.1|Acinetobacter_genomosp.|1-435	53.2	47.6	AB119105.1|Acinetobacter_genomosp.|1-435	53.2	47.6
SI_mG2_03	600	GQ343181	methyltransferase type 11	AE011788.1|Xanthomonas_axonopodis_pv._citri_str._306|4237-5232	43.1	26.3	AE014075.1|Escherichia_coli|243401-244171	42.8	20.6
SI_mG2_06	432	GQ343179	aminoglycoside N(6')-acetyltransferase type 1	AB119105.1|Acinetobacter_genomosp.|1-435	53.2	47.6	AB119105.1|Acinetobacter_genomosp.|1-435	53.2	47.6
SI_mG2_08	432	GQ343180	aminoglycoside N(6')-acetyltransferase type 1	AB119105.1|Acinetobacter_genomosp.|1-435	53.2	47.6	AB119105.1|Acinetobacter_genomosp.|1-435	53.2	47.6
SI_mG2_10	432	GQ343183	aminoglycoside N(6')-acetyltransferase type 1	AB119105.1|Acinetobacter_genomosp.|1-435	53.2	47.6	AB119105.1|Acinetobacter_genomosp.|1-435	53.2	47.6
SI_mO1_01	429	GQ343186	aminoglycoside 6'-N-acetyltransferase	AB119105.1|Acinetobacter_genomosp.|1-435	60.1	45.8	AE014613.1|Salmonella_enterica_subsp.|1582846-1583364	48.7	37.3

Table S4: Resistance genes identified using functional selections from cultured aerobic gut microbiome isolates. Gene ID is made up of the 2 letter code for the antibiotics used for the selections (e.g. AX denotes Amoxicillin, see Table S1), an i denoting cultured isolate origin of DNA insert, microbiome two letter code (e.g. G1 denotes gut microbiome 1) and a numeric identifier. For each gene identified, the most similar gene from any organism as well as the most similar gene harbored by a pathogenic isolate in GenBank was identified using tblastx. Global sequence identities at the nucleotide and amino acid level between these genes were computed using clustalW (*3*).

Gene ID	**Gene length [bp]**	**Gen-bank ID**	**Gene Annotation**	**Top hit [gbID|title|position]**	**Global DNA % ID**	**Global Protein % ID**	**Top pathogenic hit [gbID|title|position]**	**Global DNA % ID**	**Global Protein % ID**
AM_iG1_01	711	GQ342979	putative efflux pump	BA000007.2|Escherichia_coli_O157:H7_str._Sakai|218601-219311	96.9	99.6	BA000007.2|Escherichia_coli_O157:H7_str._Sakai|218601-219311	96.9	99.6
AM_iG1_03	600	GQ342980	aminoglycoside 6'-N-acetyl transferase type Ib-cr	AY458016.1|Escherichia_coli|11679-12278	100.0	100.0	AY509609.1|Pseudomonas_aeruginosa|1075-1593	99.6	98.8
AM_iG1_06	711	GQ342983	putative efflux pump	BA000007.2|Escherichia_coli_O157:H7_str._Sakai|218601-219311	94.4	98.3	BA000007.2|Escherichia_coli_O157:H7_str._Sakai|218601-219311	94.4	98.3
AM_iG1_07	600	GQ342978	aminoglycoside 6'-N-acetyl transferase type Ib-cr	AY458016.1|Escherichia_coli|11679-12278	100.0	100.0	AY458016.1|Escherichia_coli|11679-12278	100.0	100.0
AM_iG1_11	711	GQ342981	putative efflux pump	BA000007.2|Escherichia_coli_O157:H7_str._Sakai|218601-219311	96.9	99.6	BA000007.2|Escherichia_coli_O157:H7_str._Sakai|218601-219311	96.9	99.6
AM_iG1_13	447	GQ342984	putative efflux pump	CP000247.1|Escherichia_coli|219476-220186	95.7	100.0	CP000247.1|Escherichia_coli|219476-220186	95.7	100.0
AM_iG1_15	786	GQ342985	putative permease, membrane transport protein	CP000247.1|Escherichia_coli|835762-836466	99.0	100.0	CP000247.1|Escherichia_coli|835762-836466	99.0	100.0
AM_iG1_19	600	GQ342986	aminoglycoside 6'-N-acetyl transferase type Ib-cr	AY458016.1|Escherichia_coli|11679-12278	100.0	100.0	AY458016.1|Escherichia_coli|11679-12278	100.0	100.0

Gene ID	Gene length [bp]	Gen-bank ID	Gene Annotation	Top hit [gbID|title|position]	Global DNA % ID	Global Protein % ID	Top pathogenic hit [gbID|title|position]	Global DNA % ID	Global Protein % ID
AM_iG1_21	672	GQ342987	putative efflux pump	CP000247.1|Escherichia_coli|219476-220186	95.8	98.7	CP000247.1|Escherichia_coli|219476-220186	95.8	98.7
AM_iG2_01	537	GQ342992	Ribosomal RNA small subunit methyltransferase F (rRNA (cytosine-C(5)-)-methyltransferase rsmF)	CP000247.1|Escherichia_coli|1832067-1833512	99.8	99.4	CP000247.1|Escherichia_coli|1832067-1833512	99.8	99.4
AM_iG2_05	483	GQ342988	putative efflux pump	AL627266.1|Salmonella_enterica_subsp.|28279-28980	79.6	86.9	AL627266.1|Salmonella_enterica_subsp.|28279-28980	79.6	86.9
AM_iG2_07	711	GQ342989	putative efflux pump	CP000247.1|Escherichia_coli|219476-220186	100.0	100.0	CP000247.1|Escherichia_coli|219476-220186	100.0	100.0
AM_iG2_11	711	GQ342990	putative efflux pump	CP000247.1|Escherichia_coli|219476-220186	100.0	100.0	CP000247.1|Escherichia_coli|219476-220186	100.0	100.0
AM_iG2_12	578	GQ342991	putative efflux pump	CP000247.1|Escherichia_coli|219476-220186	100.0	100.0	CP000247.1|Escherichia_coli|219476-220186	100.0	100.0
AM_iG2_15	711	GQ342993	putative efflux pump	CP000247.1|Escherichia_coli|219476-220186	100.0	100.0	CP000247.1|Escherichia_coli|219476-220186	100.0	100.0
AM_iG2_17	654	GQ342994	putative efflux pump	AL627266.1|Salmonella_enterica_subsp.|28279-28980	76.7	82.5	AL627266.1|Salmonella_enterica_subsp.|28279-28980	76.7	82.5
AM_iG2_19	711	GQ342995	putative efflux pump	CP000247.1|Escherichia_coli|219476-220186	100.0	100.0	CP000247.1|Escherichia_coli|219476-220186	100.0	100.0
AX_iG1_01	861	GQ343004	TEM beta-lactamase	AY123253.3|Klebsiella_pneumoniae|24558-25418	100.0	100.0	AY123253.3|Klebsiella_pneumoniae|24558-25418	100.0	100.0
AX_iG1_04	876	GQ343005	CTX-M beta-lactamase	AY604722.1|Escherichia_coli|634-1509	100.0	100.0	AY604722.1|Escherichia_coli|634-1509	100.0	100.0
AX_iG1_06	861	GQ343006	TEM beta-lactamase	AY123253.3|Klebsiella_pneumoniae|24558-25418	100.0	100.0	AY123253.3|Klebsiella_pneumoniae|24558-25418	100.0	100.0
AX_iG1_08	864	GQ343007	CTX-M beta-lactamase	AY604722.1|Escherichia_coli|634-1509	100.0	100.0	AY604722.1|Escherichia_coli|634-1509	100.0	100.0

Gene ID	Gene length [bp]	Genbank ID	Gene Annotation	Top hit [gbID|title|position]	Global DNA % ID	Global Protein % ID	Top pathogenic hit [gbID|title|position]	Global DNA % ID	Global Protein % ID
AX_iG2_01	1263	GQ343016	beta-lactamase	CP000247.1|Escherichia_coli|4607984-4609246	100.0	100.0	CP000247.1|Escherichia_coli|4607984-4609246	100.0	100.0
AX_iG2_02	891	GQ343019	beta-lactamase	L08472.1|Bacteroides_uniformis|848-1738	99.9	100.0	AF148067.1|Escherichia_coli|214-1158	53.4	40.7
AX_iG2_03	1146	GQ343034	beta-lactamase class C	D13207.1|Citrobacter_freundii|143-1288	99.4	99.5	AY509004.1|Salmonella_enterica_subsp.|69966-71195	95.3	97.4
AX_iG2_04	1263	GQ343008	AmpC beta-lactamase	U00096.2|Escherichia_coli_str._K12_substr._MG1655|4375834-4376967	100.0	100.0	AE014075.1|Escherichia_coli|4986532-4987698	98.2	98.5
AX_iG2_05	1146	GQ343009	CMY AmpC beta-lactamase	D13207.1|Citrobacter_freundii|143-1288	99.4	99.5	D13207.1|Citrobacter_freundii|143-1288	99.4	99.5
AX_iG2_08	1167	GQ343010	AmpC beta-lactamase	U00096.2|Escherichia_coli_str._K12_substr._MG1655|4375834-4376967	100.0	100.0	AE014075.1|Escherichia_coli|4986532-4987698	98.2	98.5
AX_iG2_10	861	GQ343011	TEM beta-lactamase	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0
AX_iG2_12	861	GQ343012	TEM beta-lactamase	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0
AX_iG2_14	1263	GQ343013	AmpC beta-lactamase	U00096.2|Escherichia_coli_str._K12_substr._MG1655|4375834-4376967	100.0	100.0	AE014075.1|Escherichia_coli|4986532-4987698	98.2	98.5
AX_iG2_15	861	GQ343014	TEM beta-lactamase	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0
AX_iG2_17	1263	GQ343015	AmpC beta-lactamase	DQ092423.1|Escherichia_coli|185-1318	100.0	100.0	DQ092423.1|Escherichia_coli|185-1318	100.0	100.0
AX_iG2_19	861	GQ343017	TEM beta-lactamase	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0

Gene ID	Gene length [bp]	Gen-bank ID	Gene Annotation	Top hit [gbID|title|position]	Global DNA % ID	Global Protein % ID	Top pathogenic hit [gbID|title|position]	Global DNA % ID	Global Protein % ID
AX_iG2_21	1146	GQ343018	CMY AmpC beta-lactamase	AF349569.1|Citrobacter_freundii|1-1146	99.2	99.7	AF349569.1|Citrobacter_freundii|1-1146	99.2	99.7
AX_iG2_24	861	GQ343020	TEM beta-lactamase	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0
AX_iG2_26	699	GQ343021	AmpC beta-lactamase	CP000247.1|Escherichia_coli|4607984-4609246	100.0	100.0	CP000247.1|Escherichia_coli|4607984-4609246	100.0	100.0
AX_iG2_29	1146	GQ343022	CMY AmpC beta-lactamase	AF349569.1|Citrobacter_freundii|1-1146	99.2	99.7	AF349569.1|Citrobacter_freundii|1-1146	99.2	99.7
AX_iG2_31	1146	GQ343023	CMY AmpC beta-lactamase	AF349569.1|Citrobacter_freundii|1-1146	99.2	99.7	AF349569.1|Citrobacter_freundii|1-1146	99.2	99.7
AX_iG2_33	861	GQ343024	TEM beta-lactamase	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0
AX_iG2_35	1263	GQ343025	AmpC beta-lactamase	DQ092423.1|Escherichia_coli|185-1318	100.0	100.0	DQ092423.1|Escherichia_coli|185-1318	100.0	100.0
AX_iG2_36	1263	GQ343026	AmpC beta-lactamase	U00096.2|Escherichia_coli_str._K12_substr._MG1655|4375834-4376967	100.0	100.0	AE014075.1|Escherichia_coli|4986532-4987698	98.2	98.5
AX_iG2_38	1256	GQ343027	AmpC beta-lactamase	CP000247.1|Escherichia_coli|4607984-4609246	100.0	100.0	CP000247.1|Escherichia_coli|4607984-4609246	100.0	100.0
AX_iG2_39	720	GQ343028	CMY AmpC beta-lactamase	D13207.1|Citrobacter_freundii|143-1288	99.2	92.9	D13207.1|Citrobacter_freundii|143-1288	99.2	92.9
AX_iG2_43	1263	GQ343029	AmpC beta-lactamase	DQ092423.1|Escherichia_coli|185-1318	100.0	100.0	DQ092423.1|Escherichia_coli|185-1318	100.0	100.0
AX_iG2_44	783	GQ343030	AmpC beta-lactamase	AE014075.1|Escherichia_coli|4986532-4987698	100.0	100.0	AE014075.1|Escherichia_coli|4986532-4987698	100.0	100.0
AX_iG2_47	1263	GQ343031	AmpC beta-lactamase	U00096.2|Escherichia_coli_str._K12_substr._MG1655|4375834-4376967	100.0	100.0	AE014075.1|Escherichia_coli|4986532-4987698	98.2	98.5
AX_iG2_48	861	GQ343032	TEM beta-lactamase	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0

Gene ID	Gene length [bp]	Gen-bank ID	Gene Annotation	Top hit [gbID|title|position]	Global DNA % ID	Global Protein % ID	Top pathogenic hit [gbID|title|position]	Global DNA % ID	Global Protein % ID
AX_iG2_52	861	GQ343033	TEM beta-lactamase	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0
CA_iG1_01	861	GQ343047	TEM beta-lactamase	AY123253.3|Klebsiella_pneumoniae|24558-25418	100.0	100.0	AY123253.3|Klebsiella_pneumoniae|24558-25418	100.0	100.0
CA_iG1_02	876	GQ343048	class A extended-spectrum beta-lactamase	AY604722.1|Escherichia_coli|634-1509	100.0	100.0	AY604722.1|Escherichia_coli|634-1509	100.0	100.0
CA_iG1_03	861	GQ343049	TEM beta-lactamase	AY123253.3|Klebsiella_pneumoniae|24558-25418	100.0	100.0	AY123253.3|Klebsiella_pneumoniae|24558-25418	100.0	100.0
CA_iG1_04	855	GQ343053	TEM beta-lactamase	AY123253.3|Klebsiella_pneumoniae|24558-25418	100.0	100.0	AY123253.3|Klebsiella_pneumoniae|24558-25418	100.0	100.0
CA_iG1_05	861	GQ343054	TEM beta-lactamase	AY123253.3|Klebsiella_pneumoniae|24558-25418	100.0	100.0	AY123253.3|Klebsiella_pneumoniae|24558-25418	100.0	100.0
CA_iG1_06	1263	GQ343055	beta-lactamase	BA000007.2|Escherichia_coli_O157:H7_str._Sakai|5232091-5233224	98.9	99.2	BA000007.2|Escherichia_coli_O157:H7_str._Sakai|5232091-5233224	98.9	99.2
CA_iG1_07	861	GQ343050	TEM beta-lactamase	AY123253.3|Klebsiella_pneumoniae|24558-25418	100.0	100.0	AY123253.3|Klebsiella_pneumoniae|24558-25418	100.0	100.0
CA_iG1_09	876	GQ343051	CTX-M beta-lactamase	AY604722.1|Escherichia_coli|634-1509	100.0	100.0	AY604722.1|Escherichia_coli|634-1509	100.0	100.0
CA_iG1_11	876	GQ343052	CTX-M beta-lactamase	AY604722.1|Escherichia_coli|634-1509	100.0	100.0	AY604722.1|Escherichia_coli|634-1509	100.0	100.0
CA_iG2_01	1263	GQ343060	beta-lactamase	DQ092423.1|Escherichia_coli|185-1318	100.0	100.0	DQ092423.1|Escherichia_coli|185-1318	100.0	100.0
CA_iG2_02	861	GQ343063	TEM beta-lactamase	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0
CA_iG2_04	861	GQ343064	TEM beta-lactamase	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0

Gene ID	Gene length [bp]	Gen-bank ID	Gene Annotation	Top hit [gbID|title|position]	Global DNA % ID	Global Protein % ID	Top pathogenic hit [gbID|title|position]	Global DNA % ID	Global Protein % ID
CA_iG2_06	1146	GQ343056	CMY AmpC beta-lactamase	AF349569.1|Citrobacter_freundii|1-1146	99.2	99.7	AF349569.1|Citrobacter_freundii|1-1146	99.2	99.7
CA_iG2_08	1263	GQ343057	AmpC beta-lactamase	DQ092423.1|Escherichia_coli|185-1318	100.0	100.0	DQ092423.1|Escherichia_coli|185-1318	100.0	100.0
CA_iG2_10	1146	GQ343058	CMY AmpC beta-lactamase	AF349569.1|Citrobacter_freundii|1-1146	99.2	99.7	AF349569.1|Citrobacter_freundii|1-1146	99.2	99.7
CA_iG2_12	1146	GQ343059	CMY AmpC beta-lactamase	D13207.1|Citrobacter_freundii|143-1288	99.4	99.5	D13207.1|Citrobacter_freundii|143-1288	99.4	99.5
CA_iG2_14	1146	GQ343061	CMY AmpC beta-lactamase	AF349569.1|Citrobacter_freundii|1-1146	99.2	99.7	AF349569.1|Citrobacter_freundii|1-1146	99.2	99.7
CA_iG2_17	861	GQ343062	TEM beta-lactamase	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0
CF_iG1_01	876	GQ343069	extended-spectrum beta-lactamase CTX-M-15	AY604722.1|Escherichia_coli|634-1509	100.0	100.0	AY604722.1|Escherichia_coli|634-1509	100.0	100.0
CF_iG2_01	1224	GQ343071	beta-lactamase	X76636.1|Citrobacter_freundii|129-1274	84.8	90.0	X76636.1|Citrobacter_freundii|129-1274	84.8	90.0
CF_iG2_02	891	GQ343070	class A beta lactamase	L08472.1|Bacteroides_uniformis|848-1738	99.9	100.0	AF148067.1|Escherichia_coli|214-1158	53.4	40.7
CF_iG2_04	1029	GQ343072	CMY AmpC beta-lactamase	AF349569.1|Citrobacter_freundii|1-1146	99.1	99.7	AF349569.1|Citrobacter_freundii|1-1146	99.1	99.7
CF_iG2_06	1146	GQ343073	CMY AmpC beta-lactamase	AJ487978.2|Citrobacter_youngae|2-1145	94.3	97.4	AY960293.2|Escherichia_coli|1-1143	86.7	93.2
CF_iG2_08	1146	GQ343074	CMY AmpC beta-lactamase	D13207.1|Citrobacter_freundii|143-1288	99.4	99.5	D13207.1|Citrobacter_freundii|143-1288	99.4	99.5
CF_iG2_10	1266	GQ343075	CMY AmpC beta-lactamase	D13207.1|Citrobacter_freundii|143-1288	99.4	99.5	D13207.1|Citrobacter_freundii|143-1288	99.4	99.5
CF_iG2_12	1146	GQ343076	CMY AmpC beta-lactamase	D13207.1|Citrobacter_freundii|143-1288	99.4	99.5	D13207.1|Citrobacter_freundii|143-1288	99.4	99.5
CF_iG2_15	1146	GQ343077	CMY AmpC beta-lactamase	D13207.1|Citrobacter_freundii|143-1288	99.4	99.5	D13207.1|Citrobacter_freundii|143-1288	99.4	99.5

Gene ID	Gene length [bp]	Gen-bank ID	Gene Annotation	Top hit [gbID|title|position]	Global DNA % ID	Global Protein % ID	Top pathogenic hit [gbID|title|position]	Global DNA % ID	Global Protein % ID
CY_iG1_01	921	GQ343119	D-alanine--D-alanine ligase B	K02668.1|Escherichia_coli|146-1066	99.9	99.7	CP000034.1|Shigella_dysenteriae|122887-123807	98.9	99.0
CY_iG1_04	1218	GQ343120	aminotransferase	AE005674.1|Shigella_flexneri_2a_str._301|2415674-2416891	99.3	99.8	AE005674.1|Shigella_flexneri_2a_str._301|2415674-2416891	99.3	99.8
CY_iG1_07	921	GQ343121	cysteine and O-acetyl-L-serine efflux system	CP000247.1|Escherichia_coli|1567495-1568394	98.3	98.7	CP000247.1|Escherichia_coli|1567495-1568394	98.3	98.7
CY_iG1_09	921	GQ343122	D-alanine--D-alanine ligase B	K02668.1|Escherichia_coli|146-1066	99.9	99.7	CP000034.1|Shigella_dysenteriae|122887-123807	98.9	99.0
CY_iG1_12	1095	GQ343123	D-alanyl-alanine synthetase A	CP000247.1|Escherichia_coli|467300-468394	99.8	99.7	CP000247.1|Escherichia_coli|467300-468394	99.8	99.7
GE_iG1_01	861	GQ343134	Aminoglycoside N(3')-acetyltransferase III (ACC(3)-III)	AY458016.1|Escherichia_coli|14339-15199	100.0	100.0	AY458016.1|Escherichia_coli|14339-15199	100.0	100.0
GE_iG1_04	861	GQ343135	Aminoglycoside N(3')-acetyltransferase III (ACC(3)-III)	AY458016.1|Escherichia_coli|14339-15199	100.0	100.0	AY458016.1|Escherichia_coli|14339-15199	100.0	100.0
GE_iG1_06	867	GQ343136	Aminoglycoside N(3')-acetyltransferase III (ACC(3)-III)	AY458016.1|Escherichia_coli|14339-15199	100.0	100.0	AY458016.1|Escherichia_coli|14339-15199	100.0	100.0
MN_iG2_01	1275	GQ343143	tetracycline resistance protein tetA	AY458016.1|Escherichia_coli|5854-7053	100.0	100.0	AY458016.1|Escherichia_coli|5854-7053	100.0	100.0
MN_iG2_03	1065	GQ343144	tetracycline resistance protein tetA	AY509004.1|Salmonella_enterica_subsp.|2514-3713	100.0	100.0	AY509004.1|Salmonella_enterica_subsp.|2514-3713	100.0	100.0
OX_iG2_01	1275	GQ343149	TetA	AY333434.1|Salmonella_typhimurium|71832-73031	100.0	100.0	AY333434.1|Salmonella_typhimurium|71832-73031	100.0	100.0

Gene ID	Gene length [bp]	Gen-bank ID	Gene Annotation	Top hit [gbID|title|position]	Global DNA % ID	Global Protein % ID	Top pathogenic hit [gbID|title|position]	Global DNA % ID	Global Protein % ID
OX_iG2_02	1128	GQ343150	TetA	AY333434.1|Salmonella_typhimurium|71832-73031	100.0	100.0	AY333434.1|Salmonella_typhimurium|71832-73031	100.0	100.0
OX_iG2_03	1275	GQ343152	TetA	AY333434.1|Salmonella_typhimurium|71832-73031	100.0	100.0	AY333434.1|Salmonella_typhimurium|71832-73031	100.0	100.0
OX_iG2_04	1128	GQ343148	tetracycline resistance protein tetA	AY458016.1|Escherichia_coli|5854-7053	100.0	100.0	AY458016.1|Escherichia_coli|5854-7053	100.0	100.0
OX_iG2_06	1275	GQ343151	tetracycline resistance protein tetA	AY458016.1|Escherichia_coli|5854-7053	100.0	100.0	AY458016.1|Escherichia_coli|5854-7053	100.0	100.0
PE_iG1_01	861	GQ343159	beta-lactamase	AY123253.3|Klebsiella_pneumoniae|24558-25418	100.0	100.0	AY123253.3|Klebsiella_pneumoniae|24558-25418	100.0	100.0
PE_iG1_02	1263	GQ343155	AmpC beta-lactamase	DQ092425.1|Escherichia_coli|185-1318	100.0	100.0	DQ092425.1|Escherichia_coli|185-1318	100.0	100.0
PE_iG1_04	567	GQ343156	AmpC beta-lactamase	DQ092426.1|Escherichia_coli|185-1318	99.6	99.5	DQ092426.1|Escherichia_coli|185-1318	99.6	99.5
PE_iG1_07	1167	GQ343157	AmpC beta-lactamase	BA000007.2|Escherichia_coli_O157:H7_str._Sakai|5232091-5233224	98.9	99.2	BA000007.2|Escherichia_coli_O157:H7_str._Sakai|5232091-5233224	98.9	99.2
PE_iG1_09	876	GQ343158	CTX-M beta-lactamase	AY604722.1|Escherichia_coli|634-1509	100.0	100.0	AY604722.1|Escherichia_coli|634-1509	100.0	100.0
PE_iG2_01	861	GQ343160	beta-lactamase	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0
PE_iG2_02	861	GQ343164	beta-lactamase	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0
PE_iG2_03	1146	GQ343161	CMY AmpC beta-lactamase	AF349569.1|Citrobacter_freundii|1-1146	99.2	99.7	AF349569.1|Citrobacter_freundii|1-1146	99.2	99.7
PE_iG2_05	1263	GQ343162	AmpC beta-lactamase	DQ092423.1|Escherichia_coli|185-1318	100.0	100.0	DQ092423.1|Escherichia_coli|185-1318	100.0	100.0

Gene ID	Gene length [bp]	Genbank ID	Gene Annotation	Top hit [gbID|title|position]	Global DNA % ID	Global Protein % ID	Top pathogenic hit [gbID|title|position]	Global DNA % ID	Global Protein % ID
PE_iG2_07	1167	GQ343163	AmpC beta-lactamase	U00096.2|Escherichia_coli_str._K12_substr._MG1655|4375834-4376967	100.0	100.0	AE014075.1|Escherichia_coli|4986532-4987698	98.2	98.5
PE_iG2_09	1263	GQ343165	AmpC beta-lactamase	U00096.2|Escherichia_coli_str._K12_substr._MG1655|4375834-4376967	100.0	100.0	AE014075.1|Escherichia_coli|4986532-4987698	98.2	98.5
PE_iG2_12	387	GQ343166	TEM beta-lactamase	AF427131.1|Escherichia_coli|209-1069	100.0	100.0	AF427131.1|Escherichia_coli|209-1069	100.0	100.0
PE_iG2_13	861	GQ343167	TEM beta-lactamase	AL513383.1|Salmonella_enterica_subsp.|161911-162771	99.8	99.7	AL513383.1|Salmonella_enterica_subsp.|161911-162771	99.8	99.7
PI_iG2_01	861	GQ343171	beta-lactamase	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0
PI_iG2_03	1146	GQ343172	beta-lactamase	AF349569.1|Citrobacter_freundii|1-1146	99.2	99.7	AF349569.1|Citrobacter_freundii|1-1146	99.2	99.7
PI_iG2_04	861	GQ343175	TEM beta-lactamase	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0
PI_iG2_05	666	GQ343173	TEM beta-lactamase	AL513383.1|Salmonella_enterica_subsp.|161911-162771	99.7	99.1	AL513383.1|Salmonella_enterica_subsp.|161911-162771	99.7	99.1
PI_iG2_08	861	GQ343174	TEM beta-lactamase	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0
PI_iG2_09	861	GQ343176	TEM beta-lactamase	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0	AY883411.1|Salmonella_enterica_subsp.|167-1027	99.7	99.0
PI_iG2_10	405	GQ343177	TEM beta-lactamase	X68082.1|artificial_sequences_synthetic|1-1071	99.5	100.0	AY123253.3|Klebsiella_pneumoniae|24558-25418	100.0	100.0
SI_iG1_01	861	GQ343184	Aminoglycoside N(3')-acetyltransferase III (ACC(3)-III)	AY458016.1|Escherichia_coli|14339-15199	100.0	100.0	AY458016.1|Escherichia_coli|14339-15199	100.0	100.0

Gene ID	Gene length [bp]	Gen-bank ID	Gene Annotation	Top hit [gbID|title|position]	Global DNA % ID	Global Protein % ID	Top pathogenic hit [gbID|title|position]	Global DNA % ID	Global Protein % ID
SI_iG1_02	600	GQ343185	aminoglycoside 6'-N-acetyl transferase type Ib	AY458016.1|Escherichia_coli|11679-12278	100.0	100.0	AY458016.1|Escherichia_coli|11679-12278	100.0	100.0
TE_iG2_01	1275	GQ343187	tetracycline resistance protein tetA	AY458016.1|Escherichia_coli|5854-7053	100.0	100.0	AY458016.1|Escherichia_coli|5854-7053	100.0	100.0

Table S5: Statistics for library screening and sequencing.

Antibiotic class	**Number of clones cultured aerobes**	**Number of clones metagenomic**	**Clones sequenced cultured aerobes**	**Clones sequenced metagenomic**	**Unique genes cultured aerobes**	**Unique genes metagenomic**
Aminoglycoside	128	356	44	21	23	7
Amino acid derivative	712	2408	8	79	5	49
Amphenicols	0	12	0	12	0	2
Beta-lactam	2170	1330	240	140	78	29
Tetracyclines	272	191	52	38	9	8
Total	**3282**	**4297**	**344**	**290**	**115**	**95**

Table S6: Transposon sequences identified using functional selections from human microbiomes. Gene ID is made up of the 2 letter code for the antibiotics used for the selections (e.g. AX denotes Amoxicillin, see Table S1), an m denoting metagenomic origin of DNA insert or an i denoting cultured aerobic isolate origin of DNA insert, microbiome two letter code (e.g. O1 denotes oral microbiome 1, G2 denotes gut microbiome 2) and) and a numeric identifier. Also listed is gene ID of the resistance gene adjacent to the transposon element (Table S3 and S4). For each gene identified, the most similar gene from the GenBank was identified using tblastx. Global sequence identities at the nucleotide level between these genes were computed using clustalW (*3*).

Gene ID	**Resistance Gene ID**	**Genbank ID**	**Annotation**	**Top tblastx hit [gbID|title|position]**	**Global DNA % ID**
AM_iG1_08	AM_iG1_07	GQ342978	transposase	AY123253.3|Klebsiella_pneumoniae|18627-19331	100.0
AM_iG1_18	AM_iG1_19	GQ342986	transposase	AY123253.3|Klebsiella_pneumoniae|20489-21193	100.0
AX_iG1_02	AX_iG1_01	GQ343004	transposon Tn3 resolvase	AY123253.3|Klebsiella_pneumoniae|25601-26158	100.0
AX_iG1_05	AX_iG1_04	GQ343005	transposase, partial sequence	AY123253.3|Klebsiella_pneumoniae|26322-29327	100.0
AX_iG1_07	AX_iG1_06	GQ343006	transposon Tn3 resolvase	AY123253.3|Klebsiella_pneumoniae|25601-26158	100.0
AX_iG1_09	AX_iG1_08	GQ343007	transposase	AY604721.1|Escherichia_coli|1031-2293	100.0
AX_iG2_11	AX_iG2_10	GQ343011	transposon Tn3 resolvase, partial sequence	AY123253.3|Klebsiella_pneumoniae|25601-26158	99.7
AX_iG2_13	AX_iG2_12	GQ343012	transposase, partial sequence	AY458016.1|Escherichia_coli|8125-9639	99.7
AX_iG2_16	AX_iG2_15	GQ343014	transposase, partial sequence	CR376602.1|Aeromonas_punctata|82018-83769	99.9

Gene ID	**Resistance Gene ID**	**Genbank ID**	**Annotation**	**Top tblastx hit [gbID|title|position]**	**Global DNA % ID**
AX_iG2_20	AX_iG2_19	GQ343017	transposase, partial sequence	AY458016.1|Escherichia_coli|8125-9639	99.7
AX_iG2_25	AX_iG2_24	GQ343020	transposase, partial sequence	AY458016.1|Escherichia_coli|8125-9639	99.7
AX_iG2_34	AX_iG2_33	GQ343024	transposase, partial sequence	AJ517790.2|Aeromonas_salmonicida|9887-11642	99.6
AX_iG2_49	AX_iG2_48	GQ343032	transposon Tn3 resolvase	AY123253.3|Klebsiella_pneumoniae|25601-26158	99.6
AX_iG2_50	AX_iG2_48	GQ343032	transposase, partial sequence	AY458016.1|Escherichia_coli|8125-9639	99.7
AX_mG2_02	AX_mG2_01	GQ342996	transposase	M60018.1|Bacteroides_fragilis|138-1430	97.8
AX_mG2_06	AX_mG2_05	GQ343000	C-terminal end of transposase, IS30 family	M17124.1|Bacteroides_fragilis|8-988	100.0
CA_iG1_08	CA_iG1_07	GQ343050	transposon Tn3 resolvase	AY123253.3|Klebsiella_pneumoniae|25601-26158	100.0
CA_iG1_10	CA_iG1_09	GQ343051	transposase, partial sequence	AY123253.3|Klebsiella_pneumoniae|26322-29327	99.3
CA_iG2_03	CA_iG2_02	GQ343063	transposase protein	AY458016.1|Escherichia_coli|8125-9639	96.2
CA_iG2_05	CA_iG2_04	GQ343064	transposase protein	AY458016.1|Escherichia_coli|8125-9639	99.3
CA_iG2_18	CA_iG2_17	GQ343062	transposon Tn3 resolvase, partial sequence	AY214164.3|Escherichia_coli|99628-100185	100.0
CA_mG2_09	CA_mG2_10	GQ343036	transposase,	M60018.1|Bacteroides_fragilis|138-1430	97.8

			partial sequence		
Gene ID	**Resistance Gene ID**	**Genbank ID**	**Annotation**	**Top tblastx hit [gbID|title|position]**	**Global DNA % ID**
GE_iG1_02	GE_iG1_01	GQ343134	transposase	AY123253.3|Klebsiella_pneumoniae|20489-21193	100.0
GE_iG1_08	GE_iG1_06	GQ343136	transposase	AY123253.3|Klebsiella_pneumoniae|20489-21193	100.0
PE_iG1_10	PE_iG1_09	GQ343158	transposase, partial sequence	AY604721.1|Escherichia_coli|1031-2293	100.0
PE_iG1_11	PE_iG1_09	GQ343158	transposase, partial sequence	AY509004.1|Salmonella_enterica_subsp.|67937-73114	100.0
PE_iG2_14	PE_iG2_13	GQ343167	transposon Tn3 resolvase	AY214164.3|Escherichia_coli|99628-100185	100.0
PI_iG2_02	PI_iG2_01	GQ343171	transposase	AY458016.1|Escherichia_coli|8125-9639	99.0
PI_iG2_06	PI_iG2_05	GQ343173	transposase	AY458016.1|Escherichia_coli|8125-9639	99.7

Table S7: Shared and unique resistance genes isolated from two unrelated individuals. The number of genes derived from metagenomic libraries or from cultured aerobic isolates are categorized as unique to an individual, or shared between individuals. For each category, the subset of genes also similar (>90 % nucleotide identity) to any gene in GenBank are shown in parentheses.

	Only Individual 1[†]	Only Individual 2[†]	Shared between 1 & 2[‡]
Metagenomic	35 (5)	52 (4)	8 (8)
Cultured aerobic isolates	24 (22)	13 (10)	78 (78)

[†] Genes are defined as unique to an individual if the nucleotide sequence identity to all genes in the other individual is less than 90 %.

[‡] Genes are defined as shared between two individuals if the nucleotide sequence identity from a gene in one individual is greater than 90 % to a gene in the other individual.

Supporting references.

1. J. R. Cole *et al.*, *Nucleic Acids Res* **35**, D169 (2007).
2. R. Lutz, H. Bujard, *Nucleic Acids Res* **25**, 1203 (1997).
3. R. Chenna *et al.*, *Nucleic Acids Res* **31**, 3497 (2003).
4. G. A. Jacoby, *Antimicrob Agents Chemother* **50**, 1123 (2006).
5. G. A. Jacoby, *Clin Microbiol Rev* **22**, 161 (2009).
6. M. Barlow, B. G. Hall, *Antimicrob Agents Chemother* **46**, 1190 (2002).
7. G. A. Jacoby, K. Bush, Amino Acid Sequences for TEM, SHV and OXA Extended-Spectrum and Inhibitor Resistant ß-Lactamases, http://lahey.org/studies/
8. M. A. Larkin *et al.*, *Bioinformatics* **23**, 2947 (2007).
9. N. Saitou, M. Nei, *Mol Biol Evol* **4**, 406 (1987).
10. V. I. Enne, C. Cassar, K. Sprigings, M. J. Woodward, P. M. Bennett, *FEMS Microbiol Lett* **278**, 193 (2008).
11. Z. Leng, D. E. Riley, R. E. Berger, J. N. Krieger, M. C. Roberts, *J Antimicrob Chemother* **40**, 551 (1997).
12. S. Demaneche *et al.*, *Proc Natl Acad Sci U S A* **105**, 3957 (2008).
13. C. S. Riesenfeld, R. M. Goodman, J. Handelsman, *Environmental Microbiology* **6**, 981 (2004).