Chapter 19

Functional Metagenomics to Study Antibiotic Resistance

Manish Boolchandani, Sanket Patel, and Gautam Dantas

Abstract

The construction and screening of metagenomic expression libraries has great potential to identify novel genes and their functions. Here, we describe metagenomic library preparation from fecal DNA, screening of libraries for antibiotic resistance genes (ARGs), massively parallel DNA sequencing of the enriched DNA fragments, and a computational pipeline for high-throughput assembly and annotation of functionally selected DNA.

Key words Functional metagenomics, Antibiotic resistance genes, Resistome, Functional selections, Massively parallel DNA sequencing, High-throughput assembly, Profile HMM-based annotation, PARFuMS, Resfams

1 Introduction

The continued evolution and global spread of antibiotic resistance genes (ARGs) in pathogens has become a major clinical and public health problem [1]. The increase in number, diversity, and range of multidrug-resistant organisms limits therapeutic options to resolve infections. To effectively mitigate or counter the antibiotic resistance problem, identification and characterization of ARGs as well as their modes of transmission and mechanisms of action is crucial.

The extensive use of antibiotics has unarguably led to a widespread increase in diversity and spread of ARGs in environmental reservoirs and pathogenic bacteria [2]. However, antibiotic resistance is ancient, existing long prior to the first discovery of natural-product antibiotics by Fleming [3]. Bacteria from diverse habitats carry extensive reservoirs of ARGs, collectively termed the "resistome," which have the potential for facile transmission to pathogens [2, 4–6].

Two conventional approaches, culture-based [7] and targeted PCR-based [8], have been frequently applied to study ARGs in complex microbial communities. While cultivation in the lab is the "gold standard" for identifying bacteria with antibiotic resistance,

Peter Sass (ed.), *Antibiotics: Methods and Protocols*, Methods in Molecular Biology, vol. 1520,
DOI 10.1007/978-1-4939-6634-9_19,

a large proportion (70–80%) of bacteria are difficult to culture in the laboratory [9]. This leads to a huge under-sampling of microorganisms belonging to diverse habitats, and as such their ARGs remain unanalyzed [10]. Targeted PCR-based approaches are generally used to identify and quantify ARGs with known sequence, bypassing the need for culture. However, these methods are only able to detect previously described genes and often require cloning into expression vectors and subsequent experimentation to verify function. Furthermore, homology-based identification and characterization of ARGs in shotgun sequences of microbial communities is inherently limited to the low number of genes with high sequence similarity to previously identified genes. In addition, such in silico analyses are unable to confirm the function encoded in putative ARGs and therefore require additional experimentation.

A much more efficient and powerful technique for characterizing resistomes is functional metagenomics (Fig. 1a) [3, 11–13], wherein total community DNA is cloned into an expression vector and transformed into a susceptible (and easily cultured) indicator strain. The resulting transformant library is assayed for antibiotic resistance by plating on selective media, and surviving ARGs are sequenced and annotated. This allows analysis of 10^9–10^{10} bp of DNA in a single experiment while exploiting three key advantages over culture- or PCR-based studies [12]: (1) No need to culture recalcitrant microorganisms, (2) No prior knowledge required about ARG sequences, and (3) Resistance phenotypes are directly associated with cloned and sequenced ARGs. Recent developments in high-throughput functional metagenomics [12] allow researchers to multiplex up to 400 functional metagenomic selections on a single Illumina sequencing lane. With the custom-built tool PARFuMS (Parallel Annotation and Reassembly of Functional Metagenomic Selections, Fig. 1b) [2], researchers can now perform demultiplexing, quality-filtering, trimming, assembly of the reads into full-length metagenomic fragments, and annotation in a single automated step, substantially reducing experimental cost.

Functional metagenomics has proven to be the most efficient and powerful method for the study of antibiotic resistance mechanisms and their associated genes across a wide variety of habitats. Our lab actively uses the following protocol and pipeline to identify and characterize ARGs derived from samples collected from different environmental- and human-associated microbial communities.

2 Materials

Prepare all reagents and buffers in ultrapure water. Use nuclease-free water to set up all reactions that involve DNA. Prepare and store all reagents at room temperature unless indicated otherwise. Thoroughly follow all applicable waste disposal regulations when disposing of biological and chemical waste.

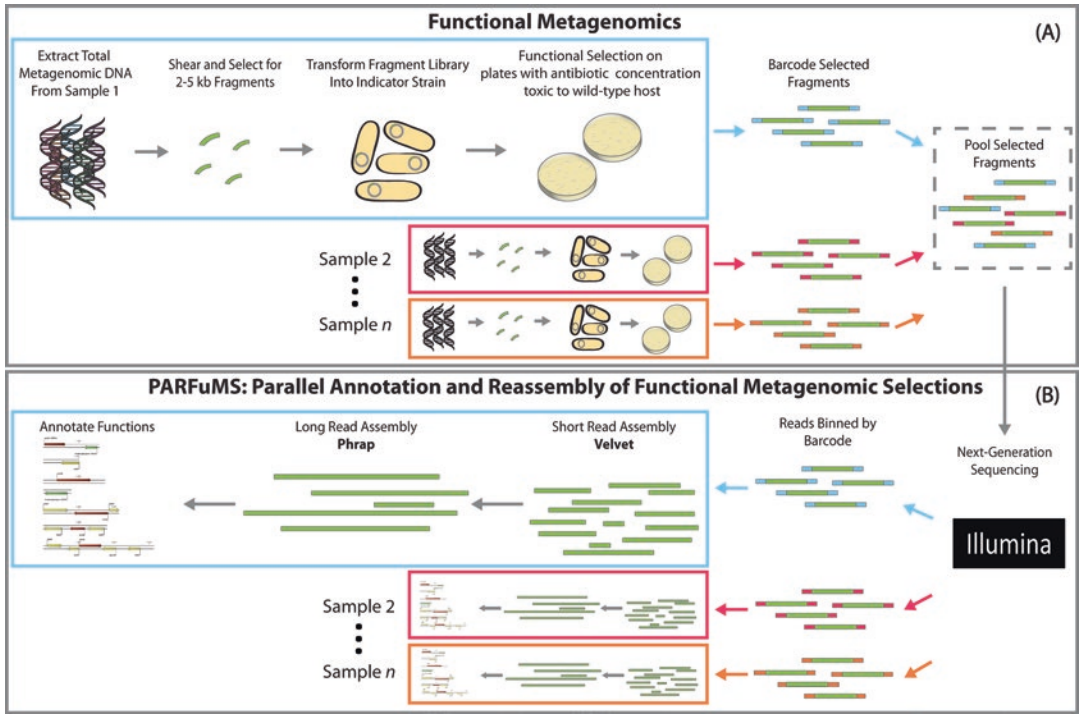

Fig. 1 Schematic representation of high-throughput functional metagenomic selection and resistome characterization. (**A**) DNA is extracted from the microbial community (e.g., feces), and 2–5 kb fragments are cloned into an expression system in an indicator strain (e.g., *Escherichia coli*). Transformants harboring resistance genes are selected using antibiotics at concentrations inhibitory to the wild-type indicator strain. Antibiotic resistance genes containing DNA fragments are PCR-amplified, sheared, bar-coded by end-ligation of oligonucleotide adapters with 5–8 bp unique sequences, and pooled in sets of up to 400 selections for next-generation sequencing. (**B**) Short (101 bp) sequencing reads from the Illumina platform are computationally binned by barcode, and assembled with computational pipeline PARFuMS, in which intermediate-length contigs from multiple rounds of assembly with the short-read assembler Velvet are assembled into full-length contigs using the long-read assembler Phrap. Contigs are annotated using custom profile Hidden Markov Models based Resfams database

2.1 DNA Extraction

1. DNA extraction buffer: 200 mM NaCl, 200 mM Tris, 20 mM EDTA in ultrapure water. Add 1.17 g of NaCl, 4 mL of 0.5 M EDTA (pH 8.0), and 20 mL of 1M Tris-HCl (pH 7.5) to ultrapure water. Make final volume to 100 mL.
2. 20% Sodium Dodecyl Sulfate (SDS).
3. 0.1 mm zirconium beads (BioSpec Products, Bartlesville, OK, USA).
4. 3 M sodium acetate, pH 5.5.
5. Phenol:Chloroform:Isoamyl alcohol (25:24:1), pH 8.0 (*see* **Note 1**).

6. Mini-Bead beater (BioSpec Products BioSpec Products, Bartlesville, OK, USA).
7. Phase-lock gel tube (PLG) (5Prime, Gaithersburg, MD, USA).
8. Molecular biology grade Isopropyl alcohol (Isopropanol).
9. Ethyl alcohol (Ethanol).
10. Tris-EDTA (TE) buffer, pH 8.0.
11. RNase A (100 mg/mL).
12. PCR Purification Kit.
13. Qubit® dsDNA BR Assay Kit (Invitrogen/ThermoFisher Scientific, Grand Island, NY, USA).
14. Qubit® dsDNA HS Assay Kit (Invitrogen/ThermoFisher Scientific, Grand Island, NY, USA).
15. Qubit® 2.0 Fluorometer (Invitrogen/ThermoFisher Scientific, Grand Island, NY, USA).

2.2 Construction of Metagenomic Library

1. 1% Low-melting point agarose gel in 0.5× Tris-Borate-EDTA (TBE) buffer with SYBR green I DNA binding dye.
2. 6× Gel loading dye.
3. pZE21 MCS-1 expression vector [14] (*see* **Note 2**).
4. High fidelity (HF) DNA Polymerase Kit (*see* **Note 3**).
5. Alkaline Phosphatase, Calf Intestinal (CIP), and corresponding 10× reaction buffer.
6. EB buffer (Qiagen, Hilden, Germany).
7. Gel Extraction Kit.
8. MinElute PCR Purification Kit (Qiagen).
9. MinElute Gel Extraction Kit (Qiagen).
10. End-It™ DNA End-Repair Kit (Epicentre, Madison, WI, USA).
11. Fast-Link™ DNA Ligation Kit (Epicentre).
12. 0.025 μm cellulose membrane (Millipore, Billerica, MA, USA).
13. 0.1 cm–gap Gene Pulser®/MicroPulser™ Electroporation Cuvettes (Bio-Rad, Hercules, CA, USA).
14. High efficiency ($\geq 4 \times 10^{10}$ cfu/μg) electrocompetent *E. coli* cells (*see* **Note 4**).
15. *Taq* DNA Polymerase with corresponding 10× reaction buffer.
16. 10 mM Deoxynucleotide (dNTP) Solution Mix.

2.3 Illumina Library Preparation and Sequencing

1. T4 DNA polymerase (New England Biolabs Inc, Ipswich, MA, USA).
2. T4 Polynucleotide Kinase (PNK) (New England Biolabs Inc, Ipswich, MA, USA).

Table 1
Commonly used antibiotics and their MICs are listed below

Antibiotics	Code	Selection concentration (μg/mL)	Antibiotic class	Antibiotic type
Aztreonam	AZ	8	β-Lactam	Synthetic
Chloramphenicol	CH	8	Amphenicol	Natural
Ciprofloxacin	CI	0.5	Fluoroquinolone	Synthetic
Colistin	CL	8	Polymyxin	Natural
Cefepime	CP	8	β-Lactam	Synthetic
Cefotaxime	CT	8	β-Lactam	Semisynthetic
Cefoxitin	CX	64	β-Lactam	Semisynthetic
d-Cycloserine	CY	32	Amino acid derivative	Natural
Ceftazidime	CZ	16	β-Lactam	Semisynthetic
Gentamicin	GE	16	Aminoglycoside	Natural
Meropenem	ME	16	β-Lactam	Semisynthetic
Penicillin	PE	128	β-Lactam	Natural
Piperacillin	PI	16	β-Lactam	Semisynthetic
Piperacillin-Tazobactam	PI-TZ	16-PI/4-TZ	β-Lactam	Semisynthetic
Tetracycline	TE	8	Tetracycline	Natural
Tigecycline	TG	2	Tetracycline	Semisynthetic
Trimethoprim	TR	8	Folate synthesis inhibitor	Synthetic
Trimethoprim-sulfamethoxazole	TR-SX	2-TR/38-SX	Folate synthesis inhibitor	Synthetic

3. T4 DNA ligase (New England Biolabs Inc, Ipswich, MA, USA).
4. 1 μM pre-annealed, Illumina barcoded sequencing adapters (*see* **Note 5**).
5. High-Fidelity PCR Master Mix.

2.4 Antibiotics and Media

1. *Antibiotics:* Purchase antibiotics in dry powdered form and prepare stock solutions as detailed in the manufacturer's MSDS. Detailed information on antibiotics is listed in Table 1 (*see* **Note 6**).
2. *Lysogeny Broth (LB) agar medium*: Add 5 g yeast extract, 10 g NaCl, 10 g tryptone, and 15 g agar in 900 mL of dH_2O. Mix well to dissolve and adjust pH to 7.5 using 1 N NaOH. Add dH_2O to 1 L. Autoclave on liquid cycle for 20 min or according to your autoclave's specification (*see* **Notes 7** and **8**).

3. *Lysogeny Broth (LB)*: Add 5 g yeast extract, 10 g NaCl, and 10 g tryptone in 900 mL of dH_2O. Mix well to dissolve and adjust pH to 7.5 using 1 N NaOH. Add dH_2O to 1 L. Autoclave on liquid cycle for 20 min or according to your autoclave's specifications (*see* **Notes** 7 and **8**).
4. *Lysogeny Broth (LB) with 15% glycerol and 50 μg/mL kanamycin:* Mix 15 mL glycerol and 85 mL autoclaved LB broth in a clean sterile flask. Mix well and filter sterilize. Add kanamycin stock solution so that the final concentration of kanamycin is 50 μg/mL.
5. *Mueller–Hinton (MH) agar:* Add 2 g beef infusion solids, 1.5 g starch, 17 g agar, and 17.5 g casein hydrolysate in 900 mL of dH_2O. Mix well to dissolve and adjust pH to 7.4 using NaOH. Add dH_2O to 1 L. Autoclave on liquid cycle for 20 min or according to your autoclave's specifications (*see* **Notes** 7 and **8**).

2.5 Instruments and Glassware

1. Thermocycler.
2. Centrifuges for 1.5 and 50 mL tubes.
3. Heat block.
4. Sonicator.
5. Electroporator.
6. Electrophoresis unit.
7. Gel trays and tank.
8. Gel imager and conversion screen.
9. Petri dish.
10. Cryo tubes.

2.6 Primers

1. pZE21 linearizing forward primer (5′ GACGGTATCGATAA GCTTGAT 3′).
2. pZE21 linearizing reverse primer (5′ GACCTCGAGGGGG GG 3′).
3. Colony PCR forward primer (5' GATACTGAGCACATC AGCAGGA 3′).
4. Colony PCR reverse primer (5' CCTGATTCTGTGGATAA CCGTA 3′).
5. Primer F1 (5′ CCGAATTCATTAAAGAGGAGAAAG 3′).
6. Primer F2 (5′ CGAATT CATTAAAGAGGAGAAAGG 3′)
7. Primer F3 (5′ GAATTCATTAAAGAG GAGAAAGGTAC 3′).
8. Primer R1 (5′ GATATCAAGCTTATCGATACCGTC 3′)
9. Primer R2 (5′ CGATATCAAGCTTATCGATACCG 3′).

10. Primer R3 (5′ TCGATATCAAGCTTATCGATACC 3′).
11. Illumina PCR Forward Primer (5′ AATGATACGGCGA CCACCGAGATCTACACTCTTTCCCTACACGACGCT CTTCCGATCT 3′).
12. Illumina PCR Reverse Primer (5′ CAAGCAGAAGACGG CATACGAGATCGGTCTCGGCATTCCTGCTG AACCGCTCTTCCGATCT 3′).

3 Methods

Carry out all procedures at room temperature unless otherwise specified. Procedures involving bacterial culture, media preparation, and transformation should be performed in a clean, sterile environment. Phenol:Chloroform:Isoamyl alcohol should be handled in the chemical hood. A PCR hood should be used to set up PCR reactions to avoid cross-contamination. Protocols may vary for different manufacturer reagents and kits. Follow manufacturer's instruction carefully and make necessary changes in protocol.

3.1 Metagenomic DNA Extraction

1. On dry ice, aliquot 50–100 mg of fecal material into a 2 mL sterile polypropylene tube (*see* **Note 9**).
2. Prepare samples for bead-beating by adding 250 μL 0.1 mm zirconium beads, 210 μL 20% SDS, 500 μL DNA extraction buffer and 500 μL phenol:chloroform:isoamyl alcohol (25:24:1; pH 8.0) to the tube containing the fecal material. Keep samples on ice for about 5 min to cool down (*see* **Note 10**).
3. Lyse microbial cells by bead-beating using a Mini-Bead beater on "homogenize" setting for a total of 4 min (bead-beating for 2 min followed by cooling the samples on ice for 2 min, and bead-beating again for another 2 min) (*see* **Note 11**).
4. Centrifuge samples at 6081 × *g* for 5 min at 4 °C.
5. Immediately prior to use, pellet Phase Lock Gel (PLG) tube at maximum speed (~16,058 × *g*) in a microcentrifuge for 20–30 s.
6. Taking care to avoid the pellet, transfer the top aqueous phase using a micropipette to a clean phase-lock gel tube and add 600 μL of phenol:chloroform:isoamyl alcohol (25:24:1; pH 8.0) to the tube. Gently mix by inversion at least ten times. Do not vortex the tubes (*see* **Note 12**).
7. Centrifuge samples at ~16,058 × *g* at room temperature for 5 min.
8. Transfer top aqueous phase from the phase-lock gel tube (~600 μL) into a clean 1.5 mL reaction tube.

9. Add 1/10 volume of 3 M sodium acetate (pH 5.5) (~60 μL) and 1 volume (~600 μL) of −20 °C isopropanol to the tube containing the aqueous phase from the previous step. Mix thoroughly by inversion.
10. Store at −20 °C for at least 2 h or overnight. DNA precipitate should be visible immediately or after incubation. *(PAUSE POINT: The samples can be stored overnight at −20 °C).*
11. After incubation, centrifuge the tube containing DNA at ~16,058 × *g* at 4 °C for 20 min. Discard supernatant carefully without disturbing the DNA pellet.
12. Wash pellet with 1 mL of 100% ethanol. Vortex to loosen the DNA pellet from the tube. Centrifuge for 5 min at ~16,058 × *g* at 4 °C. Remove ethanol without disturbing the pellet.
13. Evaporate residual ethanol by placing the sample tube on a 55 °C heat block (*see* **Note 13**).
14. Add 150 μL of TE (pH 8.0) buffer and incubate at 55 °C until DNA is completely dissolved. Gently vortex sample if needed.
15. Add 10 μL of RNase A (100 mg/mL) to the DNA sample and incubate it for 5 min at room temperature.
16. Purify DNA using the PCR Purification kit as per manufacturer's protocol.
17. Quantify purified DNA using the Qubit® dsDNA BR or HS Assay Kit and the Qubit® 2.0 Fluorometer using the manufacturer's protocols (*see* **Note 14**).

3.2 Metagenomic Library Preparation

3.2.1 Plasmid Preparation

1. Prepare the pZE21 MCS-1 expression vector for ligation by linearizing at the *Hinc*II site using inverse PCR using the blunt-end HF DNA polymerase with the following reaction conditions (*see* **Notes 2** and **3**).
   Mix the following components in a 50 μL reaction volume.

   10.0 μL 10× Polymerase reaction buffer,

   1.5 μL 10 mM dNTP mix,

   1.0 μL 50 mM $MgSO_4$,

   5.0 μL Polymerase enhancer solution,

   1.0 μL 100 pg/μL circular pZE21 DNA,

   0.75 μL 10 μM pZE21 linearizing forward primer,

   0.75 μL 10 μM pZE21 linearizing reverse primer,

   0.4 μL HF DNA Polymerase, and

   29.6 μL Nuclease-free H_2O to a final volume of 50 μL.
2. Transfer 50 μL of the above master mix to each PCR well. Run PCR on the thermocycler as follows: 95 °C for 5 min, then 35

cycles of [95 °C for 45 s, 55 °C for 45 s, 72 °C for 2.5 min], then 72 °C for 5 min (*see* **Note 15**).

3. Prepare 1 % agarose gel by adding 1 g of agarose to 100 mL of 0.5× Tris-Borate-EDTA (TBE) buffer (10× stock concentration). Heat solution until agarose is completely dissolved. Let it cool to about 65 °C by placing at room temperature, and occasionally swirl the flask to let it cool evenly. Add SYBR green I DNA binding dye (10,000× in water) to a final concentration of 1× (10 µL). Prepare gel-casting tray with comb for sample loading and pour agarose solution into casting tray while it is still liquid. Allow to solidify. Place the gel in an electrophoresis chamber, add enough 0.5× TBE buffer to cover the surface of the gel, and remove the comb.
4. Add 6× gel loading dye to the PCR-amplified DNA. Run on a 1 % low-melting-point agarose gel with SYBR green I DNA binding dye at 70 V for 120 min.
5. Excise gel slice corresponding to a ~2200 bp fragment and transfer to a clean tube. Purify DNA using the gel extraction kit. Elute DNA in 50 µL of nuclease-free molecular grade water (*see* **Notes 16** and **17**).
6. Dephosphorylate purified plasmid using calf intestinal phosphatase (CIP): For 50 µL reaction, add 40 µL of gel-purified DNA, 5 µL of CIP (10 U/µL), and 5 µL of the corresponding 10× reaction buffer. Incubate reaction mixture at 37 °C overnight and heat-inactivate reaction by incubating for 15 min at 70 °C (*see* **Note 18**).
7. Purify plasmid using the PCR purification kit as per manufacturer's protocol.
8. Quantify purified plasmid using Qubit® dsDNA HS Assay Kit and the Qubit® 2.0 Fluorometer. Store plasmid at –20 °C. Avoid multiple freeze-thaw cycles (*see* **Notes 14** and **19**).

3.2.2 Insert Preparation

1. Dilute up to 20 µg metagenomic DNA in EB buffer to a final volume of 200 µL. Shear the DNA to a size range of approximately 3 kb using a sonicator with manufacturer's recommended settings (*see* **Note 20**).
2. Add 6× gel loading dye to sheared DNA to a final concentration of 1×. Run sample through 1 % low-melting point agarose gel with SYBR green I DNA binding dye as described above (70 V for 120 min). Excise a gel slice corresponding to 2–5 kb fragment size using a clean disposable knife (*see* **Notes 17, 21** and **22**).
3. Extract metagenomic DNA from the excised gel slice using the gel extraction kit and elute in 34 µL nuclease-free water.

4. Following purification, use END-It™ DNA End Repair kit to end repair the DNA: For each 34 μL volume of size-selected metagenomic DNA, add the following:

   5 μL dNTP mix (2.5 mM),

   5 μL ATP (10 mM),

   5 μL 10× End-Repair Buffer, and

   1 μL End-Repair Enzyme Mix to a final volume of 50 μL.

   Mix gently and incubate at room temperature for 45 min. Heat-inactivate the reaction at 70 °C for 15 min.

5. Purify DNA using PCR purification kit. Elute the DNA with 30 μL of nuclease-free water.

6. Quantify purified DNA using the Qubit dsDNA HS Assay Kit and the Qubit 2.0 Fluorometer. Concentrate DNA using a vacuum concentrator or heat block at 55 °C to a final volume of 8–10 μL (*see* **Notes 14** and **23**).

3.2.3 Ligation and Dialysis

1. Perform ligation reaction using end-repaired metagenomic DNA and linearized vector using the Fast-Link™ DNA Ligation Kit using the following protocol. Maintain a 5:1 molar ratio of insert:vector for ligation. Prepare a 15 μL reaction tube with the following reagents:

   1.5 μL 10× Fast-Link buffer,

   0.75 μL 10 mM ATP,

   1 μL Fast-Link DNA ligase (2 U/ μL),

   Metagenomic DNA,

   Linearized vector, and

   Nuclease-free H_2O to a final reaction volume of 15 μL.

   Along with each set of ligations, prepare a negative control ligation reaction without any insert (i.e., metagenomic DNA) (*see* **Notes 23**–**25**).

2. Incubate reaction at room temperature overnight.

3. Heat-inactivate reaction by incubating for 15 min at 70 °C.

4. After heat inactivation, dialyze ligation reactions as follows:

5. Fill clean petri dish with 20 mL of nuclease-free water. Place a 0.025 μm cellulose membrane on top of the water so that it floats. Carefully transfer entire volume of ligated product to the membrane and close lid. Incubate for 45–60 min and carefully collect sample in clean 1.5 mL tube. Use the full reaction volume for transformation (*see* **Note 26**).

3.2.4 Electroporation, Metagenomic Library Amplification, and Quantification

1. Place a 0.1 cm-gap sterile electroporation cuvette, microcentrifuge tube, and ligated DNA on ice.
2. Thaw electrocompetent cells on ice. Mix by tapping gently (*see* **Note 4**).
3. Aliquot 25 μL of electrocompetent cells to prechilled microcentrifuge tube on ice.
4. Add entire ligation reaction volume (~15 μL) of sample to the aliquoted electrocompetent cells and stir briefly with pipet tip (*see* **Note 27**).
5. Perform electroporation using 0.1 cm cuvette with the following settings on electroporator: 10 μF, 600 Ω, and 1800 V (*see* **Note 4**).
6. Within 10 s of pulse, add 975 μL of recovery medium to cuvette and gently pipet up and down to resuspend cells. Transfer the cells and recovery medium to clean tubes.
7. Place the tube in a shaking incubator at 250 rpm for 1 h at 37 °C.
8. Repeat **steps 4–7** for negative control ligation reaction.
9. After 1 h incubation, prepare 10^{-2}, 10^{-4}, and 10^{-6} dilutions in LB-kanamycin broth of metagenomic sample libraries and negative control (no insert ligation). Also include a 10^{-1} dilution for the negative control.
10. Plate 100 μL of each dilution onto separate LB agar plates containing 50 μg/mL kanamycin. Incubate plates overnight at 37 °C. The following day, count and record the number of colonies for each plate (*see* **Note 28**).
11. Inoculate the rest of the recovered cells into 50 mL of LB broth containing 50 μg/mL kanamycin and grow overnight, shaking at 26 °C in an Erlenmeyer flask. Harvest the cells after optical density at 600 nm (OD_{600}) of the culture reaches 0.6–1.0.
12. The following day, centrifuge 50 mL overnight culture at 855 × *g* for 8 min to recover pellet. Discard liquid supernatant and resuspend pellet in 15 mL of LB broth containing 15% glycerol and 50 μg/mL kanamycin. Aliquot metagenomic library into 2 mL Cryo-tubes and store at −80 °C for subsequent screening.
13. Pick 36 random colonies from the titer plates and resuspend in 50 μL of nuclease-free water. Use this as a template for PCR reactions to estimate average insert size. Set up PCR reaction as follows: Prepare 25 μL PCR reaction for each sample with the following reagents:

    2.5 μL template DNA,

    2.5 μL 10× reaction buffer,

1.0 μL 10 mM dNTP mix,

1.0 μL 10 μM Colony PCR forward primer,

1.0 μL 10 μM Colony PCR reverse primer,

0.5 μL *Taq* DNA polymerase (5U/ μL), and

16.5 μL nuclease-free water to bring the final reaction volume to 25 μL.

Use the following thermo-cycler settings to amplify DNA: 94 °C for 10 min, 25 cycles of [94 °C for 45 s, 55 °C for 45 s, 72 °C for 5 min], and 72 °C for 10 min.

14. Run PCR amplified fragments on 1% low melting point agarose as described above. Visualize DNA fragments using a gel imager and record the size of each fragment. Calculate the average insert size and estimate total library size using the following equation (*see* **Note 29**).

$$\text{Library Size}\left(\text{GB}\right) = \frac{\left\{\text{TC} \times \left(\frac{\text{TR} - (\text{FR} + \text{NI} + \text{LS})}{\text{TR} - \text{FR}}\right) \times \text{AI}\right\}}{10^9}$$

where TC = total clones (cfu/mL of library determined from the titer plates), TR = total number of PCR reactions, FR = number of failed reactions, NI = number of colonies with no insert, determined by PCR reaction that produce ~ 300 bp band, LS = number of reactions that yield inserts less than 500 bp (estimated by subtracting 300 bp from band size), AI = average insert size after subtracting 300 bp from band size.

3.3 Screening for Antibiotic Resistance and Amplification of Antibiotic Resistance-Conferring DNA Fragments

1. Determine the minimum inhibitory concentration (MIC) of each antibiotic by plating negative control of the electrocompetent cells transformed with unmodified pZE21 on MH agar with 50 μg/mL of kanamycin (MH-Kan) and additional antibiotic. Test each batch of antibiotic using negative control before screening library. Commonly used antibiotics and related MIC are listed in Table 1.
2. Calculate the amount of library stock needed for screening. Adjust the concentration of frozen metagenomic library with LB-Kan broth such that each 100 μL aliquot of plating solution contains at least 10 times the total unique clones estimated in the library.
3. To calculate the titer of the frozen metagenomic library stock, thaw one frozen aliquot on ice, prepare 10^{-2}, 10^{-4}, and 10^{-5} dilutions, and plate on LB agar plates containing 50 μg/mL kanamycin. Incubate plates overnight at 37 °C. Count and record the number of colonies on each plate.

4. Calculate amount of library stock needed for screening using the following formula:

$$\text{Amount of library stock required} = \left[\begin{array}{l}\text{cfu / mL of library determined from the titer plates} \\ \text{post electroporation} \times 10{,}000\big] / \big[\text{titer of library} \\ \text{stock determined in previous step (cfu / mL)}\end{array}\right]$$

(*See* **Note 30**).

5. Make enough diluted library stock to screen all antibiotic plates for each library. Prepare at least 100 μL extra to set up titer plates and compensate pipetting error.
6. Plate 100 μL of diluted library on MH agar with 50 μg/mL of kanamycin (MH-Kan) and one additional antibiotic at the MIC for the negative control. Additionally, plate negative control of the electrocompetent cells transformed with unmodified pZE21 to ensure that the concentration of antibiotic used entirely inhibits the growth of clones with only pZE21 (without metagenomic insert). While setting up screening experiment, plate titers of your diluted library stock on LB-Kan plates to ensure that you plated stock with a titer in the expected range. 10^{-4} and 10^{-5} dilutions are appropriate.
7. Incubate plates for 24 h at 37 °C.
8. After incubation, inspect plates for any resistant colonies and record results.
9. Collect all resistant colonies by adding 750 μL of LB-Kan broth with 15 % glycerol to the plates and gently scraping colonies with a sterile L-shaped cell-spreader. If required, repeat this step to collect any leftover colonies.
10. Collect the slurries of functionally selected clones in 2 mL Cryo-tubes by pipette aspiration and store at −80 °C.
11. To isolate the antibiotic-resistant metagenomic inserts, thaw the stock of antibiotic-resistant slurries from the above step on ice and aliquot 300 μL of cells into a new, clean 1.5 mL reaction tube. Pellet cells by centrifuging at 16,058 × *g* for 5 min.
12. Discard supernatant. Gently wash pellet with 1 mL of nuclease-free H_2O and centrifuge at 16,058 × *g* for 5 min.
13. Discard supernatant and resuspend cells in 30 μL nuclease-free H_2O. Freeze at −80 °C for 1 h and thaw to promote cell lysis.

14. Centrifuge lysed cells at 16,058 × *g* for 2 min. Collect supernatant and use it as a template for amplification of resistance-conferring DNA fragments. For 25 μL PCR reactions, the following components are mixed together:

    2.5 μL template DNA,

    2.5 μL 10× polymerase reaction buffer,

    0.5 μL 10 mM dNTP mix,

    0.5 μL Taq DNA polymerase (5 U/μL),

    3.0 μL custom primer mix, and

    16.0 μL nuclease-free H_2O to bring the final reaction volume to 25 μL.

    The custom primer mix consists of three forward and three reverse primers. The following is the volume added from each primer stock solution (10 mM) to prepare the mix (*see* **Note 31**).

    0.5 μL of primer F1,

    0.5 μL of primer F2,

    0.5 μL of primer F3,

    0.21 μL of primer R1,

    0.43 μL of primer R2, and

    0.86 μL of primer R3

15. Perform PCR reaction using the following cycling conditions: 94 °C for 10 min, 25 cycles of [94 °C for 45 s, 55 °C for 45 s, 72 °C for 5.5 min], and 72 °C for 10 min.
16. Run PCR products on 1% agarose gel to confirm amplification.
17. Perform purification and quantification of amplified metagenomics inserts using a PCR Purification kit and Qubit dsDNA HS assay kit, respectively. Follow manufacturer's protocol for detailed instructions.

3.4 Illumina Library Preparation and Sequencing

1. Dilute 500–2000 ng of PCR-amplified metagenomics inserts from each selection to a total volume of 130 μL in EB buffer in a 96-well PCR plate and shear to 150–200 bp using a sonicator (*see* **Note 32**).
2. Purify and concentrate sheared DNA using the MinElute PCR Purification Kit and elute the DNA in 20 μL 55 °C nuclease-free H_2O. Use eluted DNA as input for Illumina library preparation.
3. End-repair sheared DNA by mixing 20 μL of eluted DNA with 2.5 μL T4 DNA ligase buffer, 1 μL 1 mM dNTPs, 0.5 μL T4 DNA polymerase, 0.5 μL T4 Polynucleotide Kinase, and 0.5 μL *Taq* DNA polymerase for a total reaction volume of 25 μL.

4. Incubate reaction mixture in a thermocycler at 25 °C for 30 min followed by 20 min at 75 °C.
5. To each end-repaired sample, add 5 μL of 1 μM pre-annealed barcoded sequencing adapters and 0.8 μL of T4 DNA ligase (*see* **Note 5**).
6. Incubate the reaction mixture on a thermocycler at 16 °C for 40 min followed by 10 min at 65 °C.
7. Run each sample on 2% agarose gel, stained with SYBR green I DNA binding dye as previously described, in 0.5× TBE buffer at 120 V for 2 h. Add 6× loading dye to DNA before loading on the gel.
8. Remove gel slice corresponding to 300–400 bp using a clean disposable knife and purify DNA using the MinElute Gel Extraction kit. Elute into 12 μL Buffer EB.
9. Set up PCR reaction to enrich purified DNA using 12.5 μL 2× Phusion® High Fidelity Master Mix and 0.5 μL of 10 μM Illumina PCR forward Primer and 0.5 μL of 10 μM Illumina PCR reverse Primer in a 25 μL reaction using 2 μL of purified DNA as a template. Amplify DNA at 98 °C for 30 s followed by 18 cycles of [98 °C for 10 s, 65 °C for 30 s, and 72 °C for 30 s] with a final extension of 5 min at 72 °C.
10. Run each sample on 2% agarose gel in 0.5× TBE, stained with SYBR green I DNA binding dye as previously described.
11. Remove gel slice in 300–400 bp size range using clean disposable knife and purify DNA using the MinElute Gel Extraction kit.
12. Perform DNA quantification using Qubit dsDNA HS assay kit.
13. Dilute each sample to 10 nM, and combine an equal volume of each 10 nM sample in one tube for Illumina HiSeq/MiSeq (*see* **Note 33**).

3.5 Computational Analysis of Sequencing Reads

We have developed a computational pipeline, PARFuMS as previously described in [2], for high-throughput assembly and annotation of resistance-conferring DNA fragments obtained from many independent functional selections. Below we have described a general workflow of the pipeline. It is important to note that the steps mentioned below include modifications that were made after the publication.

3.5.1 Demultiplexing and Preprocessing of Sequencing Reads

After sequencing, Illumina paired-end sequencing reads are demultiplexed based on barcode sequences. Each read is assigned to a sample-specific file by exact barcode matching via a mapping file. This step generates several smaller sequencing files, such that assembly and annotation for each sample can be performed in parallel.

Subsequently, reads corresponding to either expression vector or Illumina adapter sequences are removed or trimmed before proceeding to assembly (*see* **Note 34**).

3.5.2 De Novo Assembly of Short Read Sequences

In order to carry out functional characterization of the metagenomic insert library, quality-filtered short read sequences need to be assembled into longer contiguous sequences, commonly called contigs. Several assembly programs have been specifically developed to carry out this task, such as Meta-IDBA [15], Meta-Velvet [16], and InteMap [17]. We have implemented an iterative assembly approach as also described in [2]. In this method, short reads are first assembled into intermediate length contigs using three iterations of the short-read assembler Velvet [18]. Following each round of assembly, redundant contigs are collapsed to one sequence using CD-HIT [19] and chimeric sequences are removed by mapping raw reads against assembled contigs using FR-HIT [20]. The first iteration of Velvet takes all reads as an input, while in the second and third rounds, the reads not present in previously assembled contigs are utilized. The velvet-assembled, nonredundant contig set is then passed to the long-read assembler Phrap for two iterations. The first iteration assembles the Velvet output into more complete contigs that are subsequently linked together if two contigs are bridged by a sufficient number of raw paired-end reads. The final iteration of Phrap uses these linked contigs as input and provides a more complete assembly of linked contigs that are subsequently annotated using Resfams [21].

3.5.3 Annotation of Assembled Contigs with Antibiotic Resistance Functions

1. Identification of open reading frames in assembled contigs can be achieved by using gene prediction tools such as MGC [22], Metagenemark [23], MetaGenAnnotator [24], and GLIMMER-MG [25] (*see* **Note 35**).
2. Download the latest version of the Resfams database and supporting datafiles from http://dantaslab.wustl.edu/resfams (*see* **Note 36**).
3. Download the latest version of HMMER3 [26] from http://hmmer.org/; and follow instructions provided in the documentation to install HMMER3.
4. Run the hmmscan function of HMMER3 to annotate protein sequences using the Resfams database (Resfam.hmm) with the following parameters: -cut_ga, -tblout. Below is the basic usage to search against the Resfams database:

   ```
   hmmscan --cut_ga --tblout -o <output_file>
   Resfam.hmm <protein_seqs.fasta>
   ```

5. The output will list translated ORFs (antibiotic resistance proteins) present in functional selections (*see* **Note 37**).

4 Notes

1. Check pH of the packaged product before use. Some products are packaged at pH 6.7 and come with separate alkaline buffers. Add sufficient volume of alkaline buffer to Phenol: Chloroform:Isoamyl alcohol (25:24:1) (pH 6.7) to achieve pH 8.0. Store at 4 °C.
2. Other expression plasmids may be used to construct the library. While selecting another plasmid consider compatibility, copy number, and selective marker. Selecting different plasmid will also require redesigning of primers (for plasmid linearization, colony PCR, sequencing library preparation, etc.), and reestablishment of minimum inhibitory concentration for each antibiotic for *E. coli* with the new empty plasmid.
3. It is critical to use a high fidelity polymerase for efficient amplification of DNA fragments with a low error rate. The given reaction volumes and conditions for PCR in this protocol are optimized for Platinum® *Pfx* DNA Polymerase and may differ for other high fidelity polymerases. For a different polymerase, adjust PCR reaction conditions as per the manufacturer's recommendation.
4. Electroporation settings and minimum inhibitory concentration (Table 1) are optimized for *E. coli* E. cloni 10G SUPREME cells ($\geq 4 \times 10^{10}$ cfu/μg transformation efficiency) (Lucigen Corporation, Middleton, WI, USA). If using different electrocompetent cells, follow manufacturer's instructions for optimal electroporation condition. Reestablish MIC against each antibiotic with electrocompetent cells transformed with empty plasmid.
5. The adapters consist of unique 7 bp oligonucleotide sequences (barcodes) specific to each sequencing sample (e.g., one barcode used per antibiotic selection plate), facilitating the demultiplexing of barcoded reads on a sequencing run to enable separate assembly of contigs corresponding to each sample. Forward and reverse sequencing adapters are annealed by heating 1 μM mixture to 95 °C followed by slow cooling (0.1 °C per second) to a final holding temperature of 4 °C. Store preannealed barcoded adapters at −20 °C.
6. Carefully read the MSDS for each antibiotic and store antibiotics accordingly. Some antibiotics need to be stored in a refrigerator or a desiccator (or both). If antibiotics are light-sensitive, store them in a dark place. Expiration dates for solutions are earlier than for the stock powder. Antibiotic stock solutions prepared in ultrapure water need to be filter-sterilized. Antibiotic solutions prepared in organic solvents do not need filter sterilization. It is advisable to prepare stock solutions just before use and, if necessary, store at −20 °C.

7. Pre-warm the water bath to 55 °C. After removing the media from the autoclave, allow it to cool to 55 °C by placing in a pre-warmed water bath. Add the appropriate amount of the desired antibiotic to this media and pour about 20 mL of media per 10 cm sterile petri dish. When pouring plates, keep your bench area sterile and clean.
8. If your lab has premixed powder, use the amount suggested by the manufacturer.
9. Although 10–20 μg of starting DNA is required for one metagenomic library preparation, it is recommended to extract DNA sufficient for at least three metagenomic library preparations (50–60 μg). Insert yield may vary from sample to sample, and lower yield may hinder the process of library preparation. Thus, it is advisable to begin the library preparation with 10–20 μg of DNA and if insert yield is insufficient, the insert preparation step can be repeated to achieve the requisite amount. It is not recommended to perform the complete DNA extraction process again for a subset of samples as that may introduce extraction-specific biases into the study. Also, DNA preparations from too much fecal sample may have inhibitors that will interfere with downstream steps.
10. Close cap tightly to avoid any leakage during bead beating.
11. Continuous bead beating for 4 min may overheat samples. Cooling samples before and during bead beating helps prevent heat-induced DNA damage.
12. Use of phase lock gel tubes helps eliminate interphase protein contamination during phenol extraction. The organic phase and the interphase materials are effectively trapped in or below the gel, thus allowing easy removal of the top aqueous phase containing DNA by pipetting or decanting.
13. Do not over dry DNA pellet; dried DNA is sometimes difficult to dissolve.
14. DNA quantification using spectrometric methods may overestimate DNA concentration due to the presence of other compounds in the solution. Measurement of DNA concentration using a fluorimeter thus may give a more accurate concentration.
15. Do not use a high amount of circular plasmid as template. Carryover circular plasmid (template) may give a high background of self-ligated plasmid (plasmid without insert) during metagenomic library preparation.
16. Gel purification and size selection is preferred over a simple PCR cleanup to remove any primer dimers and other forms of plasmid. During gel extraction, try to avoid DNA differing from 2200 bp in size. In the final step of the gel extraction protocol, elute DNA with nuclease-free water instead of

elution buffer. Elution buffer contains salts that may interfere with the following steps of plasmid preparation. Optional: Further, purify plasmid DNA using a PCR purification kit to concentrate the DNA sample. This step may reduce the amount of CIP needed for dephosphorylation step.

17. Use conversion screen and filter while acquiring images of DNA electrophoresis samples to avoid exposing samples or users to harmful UV radiation. Use face and eye protection while working under UV lights.
18. Dephosphorylation removes the terminal 5′-phosphate group from DNA and thus suppresses self-ligation and circularization of linearized plasmid DNA. This is a very important step for high-efficiency plasmid preparation. Maintain a proper buffer ratio for optimal results.
19. Prepare enough plasmid for one project. Different batches of pZE21 prepared in this way have different ligation efficiencies, even when the same protocol has been followed for all. Quality control of each batch is required before using any new batch of plasmid. The easiest way to do this is by preparing an insert library from salmon sperm DNA and ligating it into your new vector. Once you confirm the ligation efficiency (ideally >70% of transformants should contain insert), you may start with the real ligation.
20. If using Covaris E210, we recommend the following settings: Duty cycle: 20%, Intensity: 0.1, Cycles per burst: 1,000, Treatment time: 600 s.
21. Use a separate gel box for each sample and clean the gel box before use with 10% bleach for 10 min. Wash gel boxes at least three times with dH_2O between each run to avoid cross-contamination.
22. Some commercially available molecular DNA markers (ladders) consist of various DNA fragments and may contain ARGs. Make sure that the molecular marker used during gel purification does not contain any ARGs. You can also prepare a custom molecular DNA marker by amplifying specific sizes of nonbacterial DNA.
23. 200 ng or more insert is required for ligation. If a sufficient amount of insert is not obtained, repeat insert preparation step (as mentioned in **Note 5**). It is also advisable to set up two separate ligation reactions if the total amount of insert is more than 1 μg.
24. When setting up multiple ligation reactions, it is preferable to make a master mix of ligation reagents to avoid pipetting error.
25. In this case, average insert size and plasmid size are similar (~2200 bp); thus, the mass ratio approximates the molar ratio.

26. Dialysis of samples removes salt traces. Higher salt concentrations in the DNA sample may cause arcing during electroporation. Make sure that no part of the cellulose membrane sinks into the water. If multiple samples are applied on the same membrane, keep at least 2 cm space between each sample to avoid contamination.
27. Do not vortex or pipet up and down to mix the sample. This can introduce air bubbles and warm the cells. Air bubbles in sample may cause arcing during electroporation.
28. The pZE21 expression vector has the kanamycin resistance cassette.
29. Colony PCR primers are designed such that they will amplify insert at the cloning site and approximately 150 bp of plasmid on each side. Thus, when no insert is present, these primers will yield a band of 300 bp in size. This serves as a colony PCR control. Do not count PCR reactions in any calculation that failed to amplify the 300 bp fragment. While calculating fragment size, deduct 300 bp from the size on the gel. For example, if a 1500 bp band is observed on the gel, the actual fragment size is 1200 bp. Repeat colony PCR if more than 20% of reactions failed to amplify.
30. If the concentration of the frozen library stock is higher than desired, dilute library with LB broth with 50 μg/mL kanamycin; if it is lower than the desired concentration, pellet cells by gentle centrifugation at 855 g and reconstitute in an appropriate volume of LB-Kan broth. Dilute the libraries in LB broth with 50 μg/mL kanamycin so that 100 μL plating solution contains about 10 copies of each clone. For example, if colony count following electroporation indicates that the metagenomic library contains 500,000 clones, prepare 100 μL of plating solution that contains (10*500,000) clones. To avoid underestimation of amount of library stock you need to plate, use the highest estimated colony count following electroporation and the lowest estimated titer of your library stock in the formula.
31. The staggered primer mix ensures diverse nucleotide composition during early Illumina sequencing cycles.
32. If using Covaris E210, we recommend the following settings:

    Duty cycle: 10%, Intensity: 5, Cycles per burst: 200, and Treatment time: 180 s.
33. Contact your local sequencing center for specific DNA amount required for sequencing run.
34. We use cross_match from the Phrap package with the following options to remove vector and adapter sequences:

    ```
    -gap1_only –minmatch 6 –minscore 10 –gap_init -3
    ```

35. We use the standalone version of Metagenemark with default parameters to identify open reading frames.
36. Resfams is a curated database of protein families and their associated profile HMMs that are confirmed for antibiotic resistance function using experimental methods and are organized in ontology. It is extensively used for high-throughput annotation of sequence-novel ARGs. There are two variants of the Resfams database: Resfam-core.hmm and Resfam-full.hmm. The core database of Resfams profile HMMs was trained using unique antibiotic resistance protein sequences from the Comprehensive Antibiotic Resistance Database (CARD) database, the Lactamase Engineering Database (LacED), and Jacoby and Bush's collection of curated beta-lactamase proteins. The core database of Resfams profile HMMs is supplemented with additional profile HMMs from the Pfam and TIGRFam databases to generate the full Resfams profile HMM database. The full version of the Resfams database should only be used when there is previous functional evidence of antibiotic resistance activity, such as in functional metagenomic selections.
37. The option "--cut_ga" requires that genes meet profile-specific gathering thresholds (rather than a global, more permissive, default log odds cutoff) before receiving annotation. Check HMMER3 documentation for detailed instructions on hmmscan.

Acknowledgments

The original implementation of functional metagenomics selections for interrogating resistance genes was described by J. Handelsman and colleagues in 1998. We thank A. Moore and B. Wang for protocol optimization of high-throughput versions of this method presented in this manuscript, A. Reyes and K. Forsberg for computational implementation of PARFuMS, M. Gibson for development of the Resfams database, and members of the Dantas lab for helpful general discussions and for comments on the manuscript.

This work was supported in part by the NIH Director's New Innovator Award (http://commonfund.nih.gov/newinnovator/), the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK:http://www.niddk.nih.gov/), and the National Institute of General Medical Sciences (NIGMS: http://www.nigms.nih.gov/), of the National Institutes of Health (NIH) under award numbers DP2DK098089 and R01GM099538 to G.D. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

References

1. Levy SB, Marshall B (2004) Antibacterial resistance worldwide: causes, challenges and responses. Nat Med 10. doi: 10.1038/nm1145
2. Forsberg KJ, Reyes A, Wang B, Selleck EM, Sommer MO, Dantas G (2012) The shared antibiotic resistome of soil bacteria and human pathogens. Science 337(6098):1107–1111. doi:10.1126/science.1220761
3. Perron GG, Whyte L, Turnbaugh PJ, Goordial J, Hanage WP, Dantas G, Desai MM (2015) Functional characterization of bacteria isolated from ancient arctic soil exposes diverse resistance mechanisms to modern antibiotics. PloS One 10(3). doi: 10.1371/journal.pone.0069533
4. Allen HK, Donato J, Wang HH, Cloud-Hansen KA, Davies J, Handelsman J (2010) Call of the wild: antibiotic resistance genes in natural environments. Nat Rev Microbiol 8(4):251–259. doi:10.1038/nrmicro2312
5. Sommer MOA, Church GM, Dantas G (2010) The human microbiome harbors a diverse reservoir of antibiotic resistance genes. Virulence 1(4):299–303. doi:10.4161/viru.1.4.12010
6. Monier J-M, Demanèche S, Delmont TO, Mathieu A, Vogel TM, Simonet P (2011) Metagenomic exploration of antibiotic resistance in soil. Curr Opin Microbiol 14(3):229–235. doi:10.1016/j.mib.2011.04.010
7. D'Costa VM, McGrann KM, Hughes DW, Wright GD (2006) Sampling the antibiotic resistome. Science 311(5759):374–377. doi:10.1126/science.1120800
8. Pérez-Pérez FJ, Hanson ND (2002) Detection of plasmid-mediated AmpC β-lactamase genes in clinical isolates by using multiplex PCR. J Clin Microbiol 40(6):2153–2162. doi:10.1128/jcm.40.6.2153-2162.2002
9. Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE, Relman DA (2005) Diversity of the human intestinal microbial flora. Science 308(5728):1635–1638. doi:10.1126/science.1110591
10. Daniel R (2005) The metagenomics of soil. Nat Rev Micro 3(6):470–478. doi:10.1038/nrmicro1160
11. Handelsman J, Rondon MR, Brady SF, Clardy J, Goodman RM (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chem Biol 5(10):R245–R249
12. Pehrsson EC, Forsberg KJ, Gibson MK, Ahmadi S, Dantas G (2013) Novel resistance functions uncovered using functional metagenomic investigations of resistance reservoirs. Front Microbiol 4:145. doi:10.3389/fmicb.2013.00145
13. Riesenfeld CS, Department of Plant Pathology, Microbiology Doctoral Training Program a, Goodman RM, Department of Plant Pathology, Microbiology Doctoral Training Program a, Gaylord Nelson Institute for Environmental Studies UoWM, Madison, WI 53706, USA, Handelsman J, Department of Plant Pathology, Microbiology Doctoral Training Program a (2015) Uncultured soil bacteria are a reservoir of new antibiotic resistance genes. Environ Microbiol 6(9):981–989. doi: 10.1111/j.1462-2920.2004.00664.x
14. Lutz R, Bujard H (1997) Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. Nucleic Acids Res 25(6):1203–1210
15. Peng Y, Leung HC, Yiu SM, Chin FY (2011) Meta-IDBA: a de novo assembler for metagenomic data. Bioinformatics 27(13):i94–i101. doi:10.1093/bioinformatics/btr216
16. Namiki T, Hachiya T, Tanaka H, Sakakibara Y (2012) MetaVelvet: an extension of Velvet assembler to de novo metagenome assembly from short sequence reads. Nucleic Acids Res 40(20):e155. doi:10.1093/nar/gks678
17. Lai B, Wang F, Wang X, Duan L, Zhu H (2015) InteMAP: integrated metagenomic assembly pipeline for NGS short reads. BMC Bioinformatics 16:244. doi:10.1186/s12859-015-0686-x
18. Zerbino DR (2010) Using the Velvet de novo assembler for short-read sequencing technologies. Current protocols in bioinformatics/editoral board, Andreas D Baxevanis [et al] CHAPTER:Unit-11.15. doi: 10.1002/0471250953.bi1105s31
19. Fu L, Niu B, Zhu Z, Wu S, Li W (2012) CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28(23):3150–3152. doi:10.1093/bioinformatics/bts565
20. Niu B, Zhu Z, Fu L, Wu S, Li W (2011) FR-HIT, a very fast program to recruit metagenomic reads to homologous reference genomes. Bioinformatics 27(12):1704–1705. doi:10.1093/bioinformatics/btr252
21. Gibson MK, Forsberg KJ, Dantas G (2015) Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. ISME J 9(1):207–216. doi:10.1038/ismej.2014.106

22. El Allali A, Rose JR (2013) MGC: a metagenomic gene caller. BMC Bioinformatics 14(Suppl 9):S6. doi:10.1186/1471-2105-14-S9-S6
23. Zhu W, Lomsadze A, Borodovsky M (2010) Ab initio gene identification in metagenomic sequences. Nucleic Acids Res 38(12):e132. doi:10.1093/nar/gkq275
24. Noguchi H, Taniguchi T, Itoh T (2008) MetaGeneAnnotator: detecting species-specific patterns of ribosomal binding site for precise gene prediction in anonymous prokaryotic and phage genomes. DNA Res 15(6):387–396. doi:10.1093/dnares/dsn027
25. Kelley DR, Liu B, Delcher AL, Pop M, Salzberg SL (2012) Gene prediction with Glimmer for metagenomic sequences augmented by classification and clustering. Nucleic Acids Res 40(1):e9. doi:10.1093/nar/gkr1067
26. Mistry J, Finn RD, Eddy SR, Bateman A, Punta M (2013) Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. Nucleic Acids Res 41(12):e121. doi:10.1093/nar/gkt263