Antibiotic resistance is a major threat to human health and well-being. To effectively combat this problem we need to understand the range of different resistance genes that allow bacteria to resist antibiotics. To do this the whole microbiota needs to be investigated. As most bacteria cannot be cultivated in the laboratory, the reservoir of antibiotic resistance genes in the non-cultivable majority remains relatively unexplored. Currently the only way to study antibiotic resistance in these organisms is to use metagenomic approaches. Furthermore, the only method that does not require any prior knowledge about the resistance genes is functional metagenomics, which involves expressing genes from metagenomic clones in surrogate hosts. In this review the methods and limitations of functional metagenomics to isolate new antibiotic resistance genes and the mobile genetic elements that mediate their spread are explored.

Introduction

There are numerous reviews and original papers describing the threat of antibiotic resistance e.g., see Arias and Murray. There are two general categories acquired and intrinsic resistance. Intrinsic generally refers to an existing physiological trait of an organism which all members of a particular species have; such as absence of a drug target, altered membrane structure, or a physiological state such as growth in the biofilm mode. Acquired resistance refers to a heritable change in the bacterial DNA. This can be either by the mutation of an existing gene which renders the bacteria resistant to a particular antibiotic or by acquiring a gene conferring resistance by horizontal gene transfer; which is very important in mediating the spread of antibiotic resistance. There are three general mechanisms of horizontal gene transfer; i.e., transduction, a process in which bacteriophages mediate gene transfer, conjugation, in which specialized genetic elements encode all the required proteins to transport DNA from a donor organism to a recipient, and transformation, where naked DNA is taken up by the recipient.

Acquired antibiotic resistance falls into four general categories; i.e., reduction of intracellular concentration of antibiotic via efflux or permeability barriers, alteration of the antibiotic target inactivation of the antibiotic or alternative pathways that are insensitive to antibiotics. There are many examples of all of these different types of mechanisms across the bacterial world. Efflux of the antibiotic tends to be mediated by protein pumps which actively transport noxious molecules out of the bacterial cell. Inactivating enzymes will target and destroy particular antibiotics. Alteration of the antibiotic target can be as a result of a point mutation which stops the target being susceptible to the antibiotic or by acquiring a gene encoding an enzyme which protects the target. Alternative pathways usually involve the bacterium acquiring enzymes or metabolic pathways that are insensitive to the antibiotic.

In order to gain a comprehensive understanding of the different types, mechanisms, transmission, and evolution of antibiotic resistance one needs to understand the contribution of the whole microbiota of particular ecosystems to antibiotic resistance, sometimes termed the resistome. Therefore as the majority of bacteria still cannot be cultured in the laboratory, metagenomic approaches are the best way to gain this knowledge.

Three different metagenomic approaches have been used to investigate the resistome; targeted (PCR- and/or microarray-based), sequenced-based, and functional metagenomics. In the introduction I will briefly discuss the first two methods and the last method is the subject of this review and will be discussed in more detail in the subsequent sections.

Targeted (PCR-based) metagenomics can be used to track the presence of resistance genes or gene families within and between different ecosystems. PCR primers are designed which will amplify the resistance gene of interest within metagenomes. Real-time PCR can be used to produce semi-quantitative data to determine the relative abundance of different genes in the different ecosystems. This has provided useful data about the spread of resistance genes in different environments. The major disadvantage of this approach is the fact that only known resistance genes can be tracked. Microarrays have also been prepared which can detect resistance genes in metagenomes. Microarrays also allow the detection of all known resistance genes but suffer from the same limitations as PCR based approaches in that only known resistance genes can be tracked. Furthermore both array and PCR based technologies do not give any information regarding the bacterial host of the resistance genes in metagenomes.

The advent of next generation sequencing technologies has made sequenced based metagenomics a feasible option. DNA
Table 1. Examples of host vector systems used for functional metagenomic cloning

| Vector         | Host(s)                  | Features                                                                 | Uses in functional metagenomics                                      |
|---------------|--------------------------|--------------------------------------------------------------------------|---------------------------------------------------------------------|
| pH701         | *E. coli*–*B. subtilis* shuttle vector | This is a shuttle vector containing a *colE1* origin of replication for maintenance in *E. coli* and a *B. subtilis* theta origin of replication, resulting in relatively stable replication. The plasmid is commercially available from Mo Bi Tec and contains an inducible promoter and blue/white selection for identifying clones. Insert sizes up to 18 kb could be isolated in *B. subtilis*. | Has been used for screening soil metagenomic libraries for antimicrobial activities\(^{29}\) |
| Bacterial artificial chromosome BAC | *E. coli*                  | These vectors are based on the origin of replication from the *E. coli* F factor. They can stably replicate large, up to 300 kb, inserts, allowing the isolation of operons and large genetic elements. Several BAC derivatives have been made (see below). | Used for cloning large genetic elements and resistance genes\(^{33}\) |
| pCC1BAC       | *E. coli*                  | Can be induced to high copy number and is commercially available from Invitrogen. |                                                                                  |
| pMDB14        | *E. coli*–*Pseudomonas putida* and *S. lividans* shuttle vector | A shuttle BAC that can be maintained in at least three different hosts. The library has to be made in *E. coli* then transferred by conjugation into other hosts. In the non-*E. coli* hosts the BACs are integrated into the host chromosome which may lead to expression difficulties when screening for some genes. | Allows expression of genes not expressed in *E. coli* in *S. lividans*\(^{27}\) |
| pUC family    | *E. coli*                  | Has a high copy number which can be an advantage when screening for the expression of genes which do not have their own promoter and are not toxic to the host. However can only accept small inserts. This plasmid is commercially available. | Has been used for cloning single antibiotic resistance genes\(^{35}\) |
| Fosmid pM0579 | *E. coli*                  | Can take inserts of up to 40 kb. The fosmid has been modified to express genes which are not always expressed in *E. coli* by using the phage T7 promoter to drive expression of the metagenomic DNA and making use of the lambda anti transcription termination protein N so that termination signals within the metagenomic DNA are ignored. | This has allowed the isolated of carbenicillin resistance genes which could not be isolated in the parental fosmid\(^{17}\) |

can be extracted from an ecosystem and the complete sequence obtained. The resulting, relatively short contiguous sequence read lengths obtained can then be compared with public databases to identify resistance genes.\(^{11}\) This approach has the potential to identify all the known resistance genes in a metagenome. However, it can only identify genes that have been shown by functional studies (either functional metagenomics or functional studies on antibiotic resistant cultivable bacteria) to encode for resistance genes. Furthermore sequencing studies give no information on expression of the resistance genes. Therefore data from sequenced based metagenomic studies will need to be continually reassessed as more functional information becomes available.

**Functional Metagenomics**

In a functional metagenomic project, the bacterial mix is extracted (sometimes other cells may be present e.g., human from GI tract extractions) from the ecosystem of interest, the DNA is prepared, which is subsequently shot-gun cloned into cloning vectors (Table 1 summarizes the properties of some of the host/vector systems that have been used) to generate gene libraries. *E. coli* is usually used as the cloning host although others have been used.\(^{12,14}\) The clones are then plated onto antibiotic (the concentration of the antibiotic is such that it will kill the host unless it contains a resistance gene) containing agar. The DNA sequence of the insert is then determined. The type of vector used varies depending on the size of insert required. Bacterial artificial chromosomes (BACs) can typically accept over 70 kb inserts although most libraries are of a smaller average size.\(^{15}\) The advantage of using BACs is that as well as the antibiotic resistance gene any genetic element in which it is embedded can also be isolated. Furthermore DNA sequences that give indications of the phylogenetic origins of the original host bacteria can be contained in the BAC clone. A disadvantage of using BACs is the low copy number (although this can also be an advantage as the lower copy number vectors are often more stable than higher copy number vectors) and the need for the transcription and translations signals to be efficiently recognized by the host organism. Another potential disadvantage of using BACs is that further analysis, such as mutagenesis, to prove the location of the resistance gene may be required.

Vectors which only accept small inserts can also be used for cloning (Table 1). These have the advantage of higher copy number and that vector promoters and ribosome binding sites can be fused to the cloned DNA, so the host transcription and translations systems can be used. A disadvantage is that the small size of the insert will not normally allow information about the genetic background of the resistance gene, and there will be little information available about the origin of the original host organism. However, if functional metagenomic studies are used in conjunction with sequenced based metagenomics, this disadvantage can
be overcome to some extent. Recently a system enabling high throughput analysis of functional metagenomic libraries called PARFuMS (parallel annotation and re-assembly of functional metagenomic selection) has been developed.16 In this work a small DNA insert library from metagenomic DNA was constructed and clones that could grow on antibiotics were selected. These were pooled and subject to massively parallel sequencing, de novo assembly and functional annotation. The output from this type of analysis allows the genetic elements containing resistance genes to be identified and the likely origins of the resistances genes to be determined.16 An interesting finding from this work was that almost identical genetic elements containing antibiotic resistance genes were found in both soil bacteria and diverse human pathogens, indicating that genetic elements themselves can spread into diverse environments even if the original host bacteria cannot do so.16

Functional metagenomic analysis is the only metagenomic screen that has the ability to isolate completely novel antibiotic resistance genes. However the disadvantage of the technique is that expression of the resistance gene is required. As E. coli is the still the most favored cloning host genes which do not express in this organism will be missed. Furthermore antibiotics to which E. coli is intrinsically resistant cannot be investigated by functional screens in this host. As well as these false negatives it has also been shown that some genes which do not mediate antibiotic resistance in their native host can interact in a novel way with the E. coli host to confer resistance on this surrogate host resulting in a false-positive result.11 These problems can be overcome by using hosts other than E. coli (some of these are listed in Table 1). Another approach is to use enhanced expression systems for functional metagenomic cloning in E. coli.17 To do this a system in which the phage T7 promoter was used to drive metagenomic DNA expression was constructed. Another system was constructed in which the phage lambda anti transcription termination protein N was used, resulting in termination signals within the metagenomic DNA being ignored and enhanced transcription occurring. This experimental approach resulted in the isolation of carbenicillin resistance genes from metagenomic DNA which could not be isolated using the parental fosmid vector (Table 1).

**Ecosystems Investigated by Functional Metagenomics**

A large number of different ecosystems have been investigated using functional metagenomics.30-33 These approaches have allowed the isolation of novel resistance genes from various ecosystems that would not have been isolated in any other ways.

Functional metagenomics also allows the isolation of resistance genes that could not be isolated by homology searching. For example work in our lab using functional metagenomics on the oral microbiota allowed the isolation of tet(37) which inactivates tetracycline. This gene has only limited homology to anything in the databases and is therefore unlikely to be found using any other method.18

**Functional Metagenomics Aimed at Isolating Mobile Genetic Elements**

The techniques described in the following section are not specifically designed to isolate antibiotic resistance genes. However, as gene transfer by specialized genetic elements is one of the most important mechanisms mediating the spread of antibiotic resistance it is relevant to consider functional approaches to the isolation of these.19

There is a huge range of different mobile genetic elements but they can be broken down into 3 general classes each of which is isolated by a different metagenomic approach:

Integration genetic elements not capable of independent conjugal transfer

These include the simple transposons that are capable of transposing around a genome but are not in themselves conjugal. However they can transpose onto a conjugal element or can be transferred to a new host by transformation. Methods have been developed to trap transposable elements e.g., see Dziewit et al.20 Typically these involve introducing a vector which contains a potential target site of a transposon into a host strain. Transposition into the target is detected by a phenotypic change; this has included activation of silent genes and inactivation of lethal genes. Once transposition has occurred, the vector containing the transposable element is isolated and purified. Transposon trapping followed by DNA sequencing and functional analysis could allow the isolation of novel functional mobile elements with antibiotic resistance genes. To apply this method to metagenomes, metagenomic libraries prepared in one of the host/vector systems listed in Table 1 would need to be transformed into a suitable host strain with a transposon trap vector, followed by screening for inactivation of the transposon’s target.

Integration genetic elements capable of mediating their own transfer

These elements include the conjugal transposons, also called integrative conjugal elements (ICE). They contain all the genetic material required to transfer from one host to another by conjugation. These types of elements are not capable of independent replication and in order to survive within a host they have to integrate into a host replicon.21 They almost always do this via a site-specific recombinase called an integrase. We have made use of the fact that there are regions of conservation between integrases to design oligonucleotides to use as probes to interrogate DNA sequence data in public databases to find new integrative elements and their cognate resistance genes. We have also used degenerate oligonucleotide primers to generate a set of probes with which to screen metagenomic libraries in BAC clones. This has led to the discovery of integrative genetic elements containing putative ABC transporters which could function to resist antibiotics.22

Mobilizable integrative elements

These elements, although not containing all the required functions for conjugal transfer can make use of other conjugal elements present in the same cell to mediate their transfer.21
the ICE these elements have to integrate into a host replicon to survive therefore the same techniques as described above can be used to isolate such elements from metagenomes.

Plasmids

Plasmids are capable of extra chromosomal replication and are a highly heterogeneous group of genetic elements. Some are capable of independent conjugal transfer, some are mobilizable and some just contain enough genetic information for their own replication. Plasmids can be circular or linear. A method to specifically to trap plasmids has been developed. This is called transposon-aided capture (TRACA). The method involves isolating plasmids from a metagenome, followed by treatment with a plasmid safe endonuclease to enrich the population for circular DNA molecules. Next an in vitro transposition reaction with a derivative of Tn5 containing a selectable marker and an E. coli origin of replication is performed. The reaction mix is then transformed into E. coli selecting for the resistance marker on the transposon. This technique has allowed the isolation of plasmids from different ecosystems including the human oral cavity,23 activated sludge,24 and the human gut microbiome.25 Novel plasmids containing functions that have not been found by any other studies have been found using this approach. Furthermore, a circular molecule that did not contain a plasmid origin of replication was isolated from the oral metagenome. This molecule contained an integrase gene.23 Integrase typically generate circular molecules as an intermediate in their site-specific recombination reaction. Therefore this suggests that TRACA has the potential to be able to isolate integrative elements.

Bacteriophages

Bacteriophages are bacterial viruses, and are the most abundant gene containing entities in most and possibly all ecosystems on Earth.36 Phages can either kill their bacterial host upon infection or replicate within their host without killing, a process called lysogeny. Recent, mostly metagenomic studies have shown that they are important mediators of the spread of antibiotic resistance genes.37,38 This happens when the phage genome acquires antibiotic resistance genes by transposition or recombination. A second way in which phages mediate the spread of resistance genes is by transduction. In this case the phage packaging machinery malfunctions and bacterial DNA is packaged into a phage head; if this DNA contains an antibiotic resistance gene, this can be transferred to a suitable recipient bacterium. Work from my lab has shown that phages can mediate the transfer of antibiotic resistance genes on mobile genetic elements between C. difficile strains.39

In metagenomic studies phages have been purified from different ecological niches, including waste water, the gut, and the lungs of cystic fibrosis patients.26,37,38,40 Antibiotic resistance genes in the phage metagenome were found by next generation sequence analysis. The numbers of phages and resistance genes appeared to increase after treatment with antibiotics, indicating that phages are important drivers of the spread of antibiotic resistance. One reason for the parallel increase in phage concentration and the number of antibiotic resistance genes is that antibiotics can induce lysogenic phages to enter the lytic life cycle. These phages are then capable of transferring resistance genes to suitable recipient strains.

The Future of Functional Metagenomics

Functional metagenomics is a powerful tool for the isolation and detection of novel antibiotic resistance genes and is the only method for the isolation of completely novel antibiotic resistance genes from the non-cultivable fraction of the microbiota. The combination of sequence based metagenomics with functional metagenomics in the PARFuMS system allows the rapid identification of mobile genetic elements containing antibiotic resistance genes.16

The major limitation of functional metagenomics is obtaining expression of all the resistance genes in the surrogate host. To date there have been very limited functional metagenomic studies performed in host/vector systems other than E. coli. Those systems that have been investigated include gram-negative proteobacteria and the high G/C Streptomyces lividans.27 This analysis showed that some genes could be expressed in S. lividans but not in E. coli, thus demonstrating the need to investigate hosts other than E. coli, for functional metagenomic projects. Shuttle vectors that replicate in non-E. coli hosts but which can also be transferred to E. coli to take advantage of the ease of working in the latter host have been developed and show promise (Table 1).27

Systems for the isolation of mobile genetic elements have shown that completely novel elements can be isolated, using TRACA and screening for integrase in addition TRACA has shown promise as a way of isolating integrase mediated circular intermediates. One of the limitations of TRACA is that only small circular molecules have been isolated so far and no linear plasmids have been isolated. To overcome these limitations different origins of replication can be put onto the transposon. Those would be ones that are adapted for replicating large DNA fragments such as the F-factor origin used in BAC cloning.28 Origins of replication that are active in other host systems should also be investigated as not all the circular molecules targeted by TRACA will be able to stably replicate in E. coli.

Conclusions

Functional metagenomics is the only metagenomics method available which allows the isolation of completely novel antibiotic resistance genes. The approach does have its limitations but these are almost always down to the surrogate host used for cloning not being able to faithfully express all the antibiotic resistance genes from any particular metagenome. Work on perfecting better host/vector systems seems to have lagged well behind the explosion of technologies to facilitate high throughput sequencing. We can only fully take advantage of the impressive explosion of knowledge from these next generation sequencing techniques when we are better able to express all the antibiotic resistance genes.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.
