

Review

The challenges of defining the human nasopharyngeal resistome

Lucy O'Connor^{1,*} and Robert Heyderman¹

The nasopharynx is an important microbial reservoir for the emergence and spread of antibiotic-resistant organisms. The nasopharyngeal resistome is an extensive, adaptable reservoir of antibiotic-resistance genes (ARGs) within this niche. Metagenomic sequencing decodes the genetic material of all organisms within a sample using next-generation technologies, permitting unbiased discovery of novel ARGs and associated mobile genetic elements (MGEs). The challenges of sequencing a low-biomass bacterial sample have limited exploration of the nasopharyngeal resistome. Here, we explore the current understanding of the nasopharyngeal resistome, particularly the role of MGEs in propagating antimicrobial resistance (AMR), explore the advantages and limitations of metagenomic sequencing technologies and bioinformatic pipelines for nasopharyngeal resistome analysis, and highlight the key outstanding questions for future research.

Introduction

AMR has become a worrying threat to global public health, in part due to widespread use and misuse of antibiotics. Resistance is now detected to all antibiotics, including those that form the last line of defence against multidrug-resistant infections [1]. It is predicted that deaths due to AMR will rise to 10 million per annum by the year 2050 [2].

The **resistome** (see [Glossary](#)) is an extensive and adaptable reservoir of ARGs that exist within a microbial niche [3]. The human resistome is largely derived from harmless colonising bacteria; bacterial pathogens account for just a small fraction of ARGs within these communities [4]. The resistome is dominated by low-risk ARGs that are dormant, expressed at low levels, or are less able to transfer between bacteria, including precursor genes, which have the potential to generate resistance under antibiotic selective pressure, and cryptic resistance genes, which are embedded in bacterial chromosomes [3]. However, some colonising bacteria harbour ARGs colocalised with MGEs that can transfer to bacterial pathogens, causing antibiotic-resistant, life-threatening, invasive disease [5,6].

The nasopharynx is an important microbial reservoir for the emergence and spread of AMR organisms that cause pneumonia, meningitis, and sepsis, particularly in children aged under 5 years [7–9]. Key pathogens include *Staphylococcus aureus*, *Streptococcus pyogenes*, *Streptococcus pneumoniae*, and *Haemophilus influenzae*; the latter two have sufficient associated mortality to warrant vaccination programs, which in turn have promoted epidemiological shifts in carriage [8]. While there are multiple studies demonstrating the impact of antibiotics on AMR amongst individual bacterial species [9–11], analysis of the overall nasopharyngeal resistome remains challenging. There has been comparatively more investigation of the lower respiratory tract resistome, a separate ecological niche with similar sequencing challenges. Here, the search for targeted antibiotic strategies based upon individual resistome analysis has been driven by the

Highlights

The nasopharynx is an important, understudied, reservoir for emergence and spread of antibiotic-resistant bacteria that cause invasive disease, including pneumonia, meningitis, and sepsis, a leading cause of mortality in children under 5 years of age worldwide.

Metagenomic sequencing techniques provide a culture-unbiased, high-resolution tool for characterising nasopharyngeal resistome composition, with the ability to capture microbial taxonomy, function, and resistance from a single specimen.

Challenges in nasopharyngeal resistome characterisation include low sample biomass, high burden of human host DNA, the need to functionally validate novel ARGs, varying bioinformatic approaches, and limited computational resources.

¹Mucosal Pathogens Research Unit, Division of Infection and Immunity, University College London, London, UK

*Correspondence: lucy.oconnor@ucl.ac.uk (L. O'Connor).



abundant AMR in patients with chronic respiratory disease (e.g., cystic fibrosis), who have a high respiratory bacterial burden and prolonged antibiotic exposure [12].

Despite advances in culture-based and rapid-identification techniques, sometimes referred to as ‘culturomics’ [13,14], there remain uncultivable bacteria in the mucosal microbial community [15]. Next-generation **metagenomic sequencing** decodes the genetic material of all organisms within a sample [4], allowing more unbiased characterisation of bacterial diversity, ARGs, and MGEs [1,4,16,17]. Technological advances have improved affordability and throughput capacity [1], such that studies exploring the use of metagenomic sequencing for surveillance of emerging AMR in agricultural environments are now possible [18].

However, there are several challenges to effectively using metagenomics to study the nasopharyngeal resistome and its contribution to invasive infections. Nasopharyngeal swabs typically yield a low bacterial biomass with sequencing outputs that are dominated by human host DNA, thus reducing sensitivity of bacterial DNA detection [19–21]. Relating detected ARGs to phenotypic resistance is problematic, due to a lack of information about the genomic context of these resistance determinants [22].

In this review, we explore current understanding of the nasopharyngeal resistome, particularly the role of MGEs in propagating AMR; we examine the advantages and limitations of current metagenomic sequencing technologies and bioinformatic pipelines for nasopharyngeal resistome analysis and highlight outstanding questions for future research.

The resistome

Origins of the resistome

Genes with AMR potential are ubiquitous in the natural environment, driven by the evolutionary processes that have enabled bacteria to survive in proximity to other antibiotic-producing microbes within the same microbial niche [3,23]. Therefore environmental bacteria provide an inexhaustible genomic reservoir for the development and dissemination of AMR to human pathogens [4,6]. Consequently, environmental resistome surveillance studies highlight potential AMR evolution, but not all ARGs pose a risk to public health [6,24]. Nonetheless, resistome analysis provides a broad understanding of the complex evolutionary processes in AMR, identifying the genetic origin and mechanisms of dissemination. Resistome analysis may also allow us to alter our use of antibiotics to exploit synergistic combinations, enhance the bactericidal effect of existing antibiotics, and reduce selection pressure [25,26]. For example, the synergy between mecillinam and cefotaxime against extended-spectrum β -lactamase *CTX-M-15*-mediated resistance [27], where mecillinam promotes a resistance mutation that confers increased susceptibility to cefotaxime, suppressing further resistance evolution.

Structure of the resistome

The resistome consists of intrinsic and acquired (extrinsic) ARGs (Figure 1). Intrinsic ARGs are ancient and predate therapeutic use of antibiotics [28]. These are typically involved in basic bacterial physiological processes and generate resistance to an entire antibiotic class [6,29]. They are often universal within a bacterial species and largely independent of antibiotic selective pressure [29]. For example, intrinsic genes encode the outer membrane of Gram-negative bacteria, which is impermeable to glycopeptide antibiotics, rendering them resistant to glycopeptides, whereas Gram-positive bacteria do not have an outer membrane and are therefore generally susceptible [30].

Many intrinsic ARGs are inactive evolutionary precursors with inducible resistance function under antibiotic selective pressure, representing a large reservoir of AMR [3]. These cryptic resistance

Glossary

Binning: the clustering of contigs that originate from the same species.

Contig: a set of overlapping DNA segments that together contain a contiguous genomic region.

Fitness cost: reduced competitive ability of an organism to replicate and survive.

Horizontal gene transfer (HGT): the nonhereditary transmission of genetic material between bacterial organisms by transformation, transduction, or conjugation.

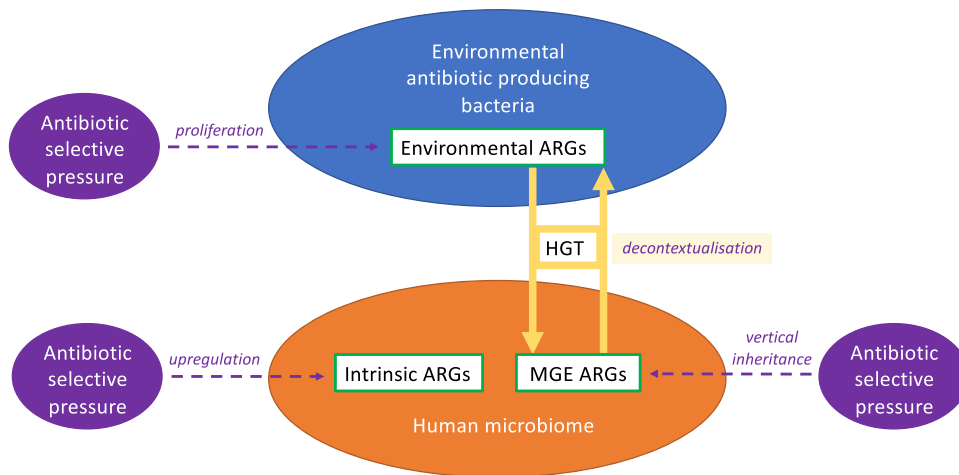
Integron: a genetic element that can capture and disseminate genes using a site-specific recombination mechanism.

K-mer: a nucleotide sequence of a certain length, for example, a dinucleotide is a k-mer where $k = 2$.

Metagenomic sequencing: a process of decoding the genetic material of all organisms within a sample.

Microbiome: the collective genetic material of the microbial community or commensal, symbiotic and pathogenic organisms within one ecological area.

Resistome: reservoir of antibiotic-resistance genes within a microbial niche.



Trends in Microbiology

Figure 1. The role of antibiotic selective pressure on the evolution of the resistome. The purple italicised text highlights the role of antibiotic selective pressure on each part of the resistome, as indicated by the purple dashed arrows. The yellow arrows illustrate the route of horizontal gene transfer (HGT) between environmental and human **microbiomes**. Abbreviations: ARGs, antibiotic-resistance genes; MGE, mobile genetic element.

genes are often embedded in bacterial chromosomes, expressed at a low level or not at all, and therefore may not result in AMR [1], for example, multidrug efflux pump gene *hmrM*, which is chromosomally encoded in *H. influenzae* [31].

Extrinsic ARGs can be acquired by *de novo* mutation or **horizontal gene transfer (HGT)**. Bacteria have a remarkable ability to mobilise genetic material within genomes by transposition, and between genomes via conjugation, interorganism contact with DNA exchange; or transformation, the direct uptake of naked DNA, on plasmids or as linear DNA released from dead cells; or phage-mediated transduction [3].

MGEs in AMR development

ARGs in MGEs, the leading drivers of AMR evolution [32], pose a greater risk to human health than chromosomally embedded ARGs [5,6]. MGEs include transposons; entire resistance gene cassettes sandwiched between **integrons**; and integrative and conjugative elements (ICEs), which often contain virulence factors (e.g., antitoxin systems) promoting their vertical inheritance within the bacteria [4,33].

Although MGE-mediated genetic exchange is more common between less taxonomically related bacteria [4], a gene with AMR potential that is acquired through HGT is removed from its original regulatory and metabolic networks. This decontextualisation alters gene expression, and therefore function, leading to constitutive gene expression or upregulation, which can generate AMR [24]. Thus, genes with a ‘housekeeping’ role in their original bacterial setting (e.g., cell wall structural proteins) can become potent AMR determinants.

β -Lactam antibiotics are particularly vulnerable to MGE-mediated AMR. A study of the nasal resistome in infants with cystic fibrosis identified MGEs for over 30% of β -lactamase genes [34]. Long-term macrolide therapy results in an increase in transmissible plasmid-encoded *ermB* amongst viridans streptococci [35], a group of nasopharyngeal commensals that are known to be a source of MGEs for more pathogenic pneumococcal species [36].

Antibiotic selective pressure

When ARGs acquired by HGT provide a survival advantage, antibiotic use promotes resistome expansion, ensuring their persistence within the resistome [4,37,38]. Even subtherapeutic antibiotic concentrations, such as those found in environmental sources due to agricultural use, can modulate gene transcription, ensuring ARG persistence within a genome [3]. However, in the absence of antibiotic selective pressure, HGT events can be lost by genetic drift [37].

The nasopharyngeal resistome

There are limited studies of the nasopharyngeal resistome (Table 1), characterised by relatively small sample sizes and single sampling time points, preventing assessment of resistome evolution. Differing study populations, DNA extraction methods, and sequencing techniques make it challenging to compare their results.

Unfortunately, most of these studies do not analyse the nasopharyngeal microbiome, focussing instead on resistance determinants or streptococcal species, particularly *S. pneumoniae* [10,11,22,34]. *S. pneumoniae* readily acquires ARGs by HGT on MGEs, and from environmental DNA by transformation [39], particularly in the context of multiple serotype carriage, which enhances the opportunities for HGT [22]. Lack of microbiome analysis precludes identifying whether the ARGs identified are present in pathogenic bacteria, which is key to assessing their risk to human health. Furthermore, most studies published to date do not present an unbiased sample of the resistome; they use targeted techniques for sequencing macrolide resistance genes [10,11,40], or enrichment for streptococci prior to DNA extraction [22,41]. This creates the incorrect impression that the nasopharyngeal microbiome and resistome are dominated by *S. pneumoniae* [42]. Other pathogens in the nasopharynx exhibit AMR, including *H. influenzae* and *S. aureus* [43]. Importantly, microbiome diversity may facilitate genetic exchange of ARGs within a resistance reservoir [42].

Nonetheless, there are some interesting themes from these studies. Firstly, whilst the prevalence of ARG carriage in the nasopharynx is high, there are varying estimates of the prevalence of phenotypic resistance. Functional metagenomics identified a resistance phenotype to β -lactam antibiotics, tetracyclines, trimethoprim-sulfamethoxazole, and chloramphenicol in 53% of nasopharyngeal samples [34], whilst a study of pneumococcal carriage showed that just 52% of isolates carrying ARGs demonstrated phenotypic resistance to penicillins, and only 39% of macrolide-resistant isolates carried ARGs for this phenotype: *msrD*, *mefA*, or *ermB* [22]. This demonstrates the complexity of conferring phenotype based on ARG carriage. Secondly, whilst most ARGs are associated with streptococcal species, β -lactamase genes are predominantly found on non-streptococcal species, which may provide a source for HGT between species [41]. For example, functional metagenomics has identified β -lactamase genes *bla*_{TEM-1} and *bla*_{BRO-type} on *H. influenzae* and *Moraxella catarrhalis*, and *bla*_Z on *S. aureus* and coagulase-negative staphylococci [34].

Thirdly, in keeping with other bacterial niches, colonisation of the nasopharynx in neonates includes ARG originating from environmental sources [40,44]. This is particularly relevant in hospitalised neonates, who can acquire nosocomial colonisers that contain ARGs on MGEs, with the potential for HGT to nasopharyngeal pathogens. For example, the *bro* β -lactamase gene is a common MGE in *M. catarrhalis*, with hypothesised origins in *Psychrobacter*, *Acinetobacter* and *Neisseria* species [45].

It is important to identify which ARGs produce a resistance phenotype as not all ARGs will be functional. Predictions of resistance phenotype based on sequence data alone are prone to false-positive bias [1]. For example, low concordance has been found between penicillin

Table 1. Summary of nasopharyngeal resistome studies

Year of study	Sample type Healthy versus diseased participants	DNA extraction and sequencing methods	Results/conclusions	Limitations	Refs
2020	Nasopharyngeal swabs Healthy infants over first year of life	<ul style="list-style-type: none"> - DNA extracted using enzymatic lysis. - Shotgun sequencing performed. - ARG-ANNOT database used. - BLAST pairwise sequence alignment. 	<ul style="list-style-type: none"> - 48 ARGs, plus resistance-associated mutations were detected; but demonstrated poor correlation with phenotypic resistance. - ARGs included macrolide-lincosamide-streptogramin B resistance (<i>msrD</i>, <i>mefA</i> and <i>ermB</i>), and tetracycline resistance (<i>tetM</i>) genes. - ARGs identified in pneumococcal resistome were <i>msrD</i>, <i>mefA</i>, <i>ermB</i> and <i>tetM</i>. 	<ul style="list-style-type: none"> - Enrichment for streptococci using Todd-Hewitt broth performed prior to DNA extraction. - No assessment of nasopharyngeal microbiome outside of pneumococci. - Sample selection not representative of study population. 	[22]
2020	Nasopharyngeal swabs Healthy infants	<ul style="list-style-type: none"> - DNA extracted using enzymatic lysis. - No host DNA depletion. - Shotgun sequencing. - ARG-ANNOT database used. - BLAST pairwise sequence alignment. 	<ul style="list-style-type: none"> - ARG were detected in 64% of samples. - Predominantly encoded resistance mechanisms to β-lactam (52%), macrolide-lincosamide-streptogramin (17%), and tetracycline antibiotics (12%). - Different type and abundance of ARG found in streptococcal and non-streptococcal contigs. - Predominantly β-lactamase genes detected from non-streptococcal reads including <i>blaOXA-60</i>, <i>blaOXA-22</i> and <i>blaBRO-1</i>. 	<ul style="list-style-type: none"> - Enrichment for streptococci using Todd-Hewitt broth performed prior to DNA extraction. - Low stringency criteria of $\geq 90\%$. - Enrichment of streptococci altered resistome composition. - ARG-ANNOT database excludes chromosomal mutations, limiting ARG detection (including PBP gene mutations and cotrimoxazole resistance) - Short-read sequencing only limits ability to assess coexistence of ARGs and MGEs. 	[41]
2019	Nasopharyngeal swabs Infants with cystic fibrosis in first year of life	<ul style="list-style-type: none"> - Functional metagenomic approach using <i>E. coli</i> as heterologous host. - DNA extracted using enzymatic lysis. - No host DNA depletion. - Long read sequencing. - Used NCBI nucleotide database, CARD, ARDB and ResFinder. - BLAST pairwise alignment. 	<ul style="list-style-type: none"> - Identified at least one resistant phenotype in 53% of samples. - Sequencing revealed a large reservoir of ARGs and MGEs within antibiotic naive samples, including ESBLs. - Antibiotic treatment may not be major factor in early resistome development. 	<ul style="list-style-type: none"> - Sensitivity of functional metagenomics limited in low-biomass samples. - Multiple displacement amplification performed. - Unable to estimate functional metagenomic library size. - <i>E. coli</i> is a poor host for discovery of ARGs from Gram-positive organisms. - Low stringency criteria of $\geq 85\%$. 	[34]
2021	Nasopharyngeal swabs Low-birthweight preterm neonates	<ul style="list-style-type: none"> - Total DNA extracted for qPCR with a commercial kit, with adjustments for Gram-positive bacteria. - No host DNA depletion. - Targeted sequencing of 16S rRNA gene and for ARGs on microbial DNA qPCR array for antibiotic resistance genes. 	<ul style="list-style-type: none"> - Absence of ARGs in nasal resistome at birth, but ARGs detected after NICU admission and increased with prolonged hospitalisation. - Most ARGs in the respiratory mucosa also detected in NICU environment, including β-lactam. - (SHV), quinolone (<i>QnrS</i>), and macrolide resistance genes (<i>ermA</i>, <i>ermB</i>, <i>ermC</i>, <i>mefA</i>, <i>msrA</i>). 	<ul style="list-style-type: none"> - Not a true metagenomic study; used two targeted sequencing techniques. - Small study size; 55 samples from 30 neonates over different time points. - No functional assessment of ARGs detected. 	[40]
2020	Nasopharyngeal swabs Healthy children from an intrapartum azithromycin randomised clinical trial	<ul style="list-style-type: none"> - DNA extraction with enzymatic lysis. - No host DNA depletion. - Targeted enrichment for macrolide resistance genes <i>msr(A)</i>, <i>ermC</i> with PCR and gel electrophoresis. - No ARG database or alignment tool used. 	<ul style="list-style-type: none"> - Intrapartum azithromycin increases carriage of macrolide-resistance genes (<i>msr(A)</i> and <i>ermC</i>) in the infant in the first month of life; this increase is not sustained at 1 year. - Increase in <i>ermC</i> and <i>msr(A)</i> prevalence at day 28 was associated with increased macrolide resistance in study population. 	<ul style="list-style-type: none"> - Not a true metagenomic study. - Small study size. - Limited time points. - No assessment of nasopharyngeal microbiome, and whether this affected prevalence of macrolide ARGs. 	[10]
2018	Nasopharyngeal swabs Healthy children aged 0–60 months	<ul style="list-style-type: none"> - DNA extracted using enzymatic lysis. - Targeted PCR amplification of <i>lytA</i>, <i>ermB</i> and <i>mefA/E</i> regions. - No host DNA depletion. - No ARG database or alignment tool used. - E-test strips used to detect macrolide resistance in <i>S. pneumoniae</i> isolates. 	<ul style="list-style-type: none"> - 20% median prevalence macrolide resistance determinants among preschool children. - Macrolide resistance determinants were more prevalent in biannually treated clusters by 24 months: 60% median prevalence versus 40% ($P < 0.001$). - Clusters randomised to less frequent azithromycin MDA had a significantly lower prevalence of macrolide resistance determinants. 	<ul style="list-style-type: none"> - Not a metagenomic study; limited to two macrolide resistance determinants. - Study randomised clusters to different dose and frequency of azithromycin; difficult to determine their separate effect on the resistome. - No assessment of nasopharyngeal microbiome, and whether this affected prevalence of macrolide ARGs. 	[11]

susceptibility and ARG presence among strains of *S. pneumoniae* [22]. The genomic context of an ARG affects its ability to function as a resistance determinant. Proximity to promoter regions will influence ARG expression, colocation with an MGE will enable it to mobilise between species (for example, streptococcal species within the nasopharyngeal microbiome) and host compatibility will determine whether it functions as a resistance determinant [6,46]. However, it is also important to recognise the fallibility of using a Gram-negative heterologous host, favoured in functional metagenomic studies for ease of culture and vector transformation, to examine phenotypic function in the resistome of a niche dominated by Gram-positive organisms [34,47,48].

Although there are many pressures that can influence the nasopharyngeal microbiome, and therefore its resistome – for example age, season and environment [8] – a key theme amongst these studies is the transient effect of antibiotic selective pressure on the nasopharyngeal resistome. Sputum samples originating from the lung show a core respiratory resistome, dominated by macrolide resistance, which remains stable despite disease state and antibiotic pressure [49,50]. In contrast, the nasopharyngeal resistome appears to change with antibiotic pressure. Several studies show an increase in macrolide-associated ARGs, particularly amongst streptococcal and staphylococcal species, following prolonged macrolide exposure [10,11,35]. This change is transient [51,52], suggesting that wild-type susceptible bacteria reclaim the nasopharyngeal niche in the absence of continued antibiotic selective pressure, potentially due to the **fitness cost** of resistance carriage [52]. Some ARGs (*mefA/E*) persist longer than others [11] and are predominantly associated with streptococcal carriage [10,53]. Other ARGs (*msr(A)* and *ermC*) are carried by MGEs, indicating a potential for HGT [54], particularly among staphylococcal and enterococcal species [55].

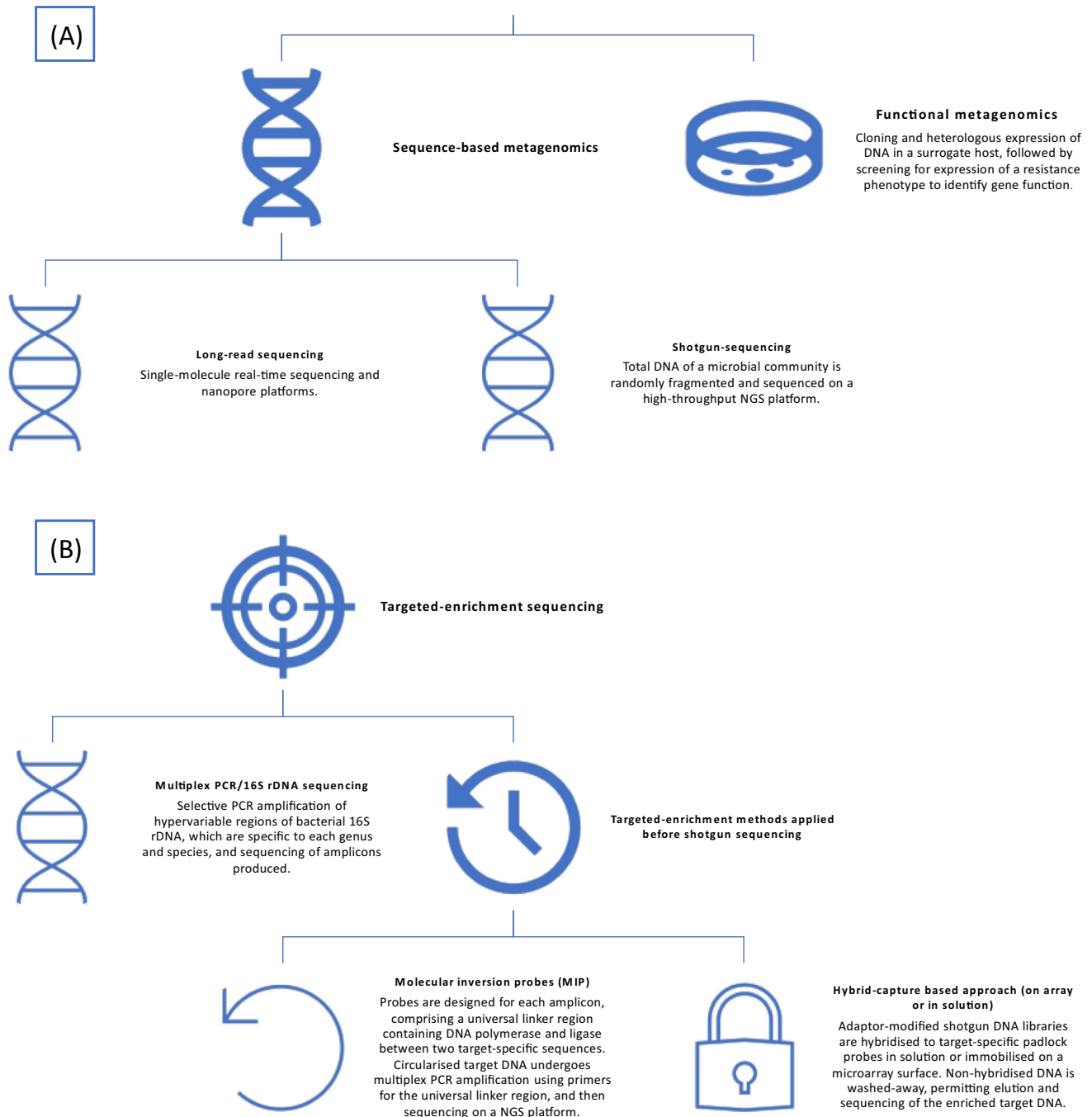
Metagenomic sequencing technologies

Targeted and shotgun metagenomics

Metagenomic sequencing for resistome analysis involves extraction and sequencing all DNA from a biological sample, processing the resulting reads in a bioinformatic pipeline, and annotating reads based on homology to catalogued sequences in ARG databases. There are two main approaches to metagenomic sequencing [4]: functional and sequence-based, as summarised in Figure 2. These approaches are often preceded by depletion of host DNA, the methods for which are detailed elsewhere [20,56]. Sequence-based approaches can be further categorised into targeted-enrichment and shotgun techniques, although only the latter is a true metagenomic method [19].

Targeted-enrichment techniques selectively capture and enrich genomic regions of interest from a sample prior to sequencing [57]. This increases the proportion of reads of interest, providing better sensitivity for the target region, but limits sequencing breadth and does not represent the complete sample metagenome [19,57]. Targeted techniques cannot be used for novel ARG discovery [23], and are limited by the expense of generating target-specific probes and the time required to amplify genomic targets [57].

Shotgun sequencing is rapid, relatively unbiased and as a truly metagenomic technique can be used for novel ARG discovery [58]. There are challenges to using sequence-based approaches on clinical samples to derive the nasopharyngeal resistome (Table 2). As mentioned, nasopharyngeal swabs yield a low bacterial biomass, thus contaminants and host DNA can conceal bacterial signals, preventing pathogen and ARG detection [19,46]. Clinical sample storage, biological variation in GC content (affecting DNA thermal stability) and extraction methods can all introduce bias before sequencing [46]. Shotgun sequencing techniques produce short, fragmented reads



Trends in Microbiology

Figure 2. Illustrating the two major approaches to metagenomic sequencing: functional and sequence-based (A). Targeted-enrichment techniques (B) are often used prior to shotgun sequencing; multiplex PCR is a targeted-enrichment technique used as a proxy for metagenomic sequencing. Abbreviation: NGS, next-generation sequencing.

that lack genomic context [58]. Direct analysis of these short reads can lead to incorrect mapping based on sequence homology and false-positive identification of ARGs [59]. Short reads require assembly into longer **contigs** to improve contextualisation of ARGs within the genome [23].

There are also universal challenges to sequencing the resistome that apply to the nasopharyngeal niche, for example functional validation of novel ARGs [1].

Long-read sequencing to provide genomic context for resistome analysis

Long-read sequencing could prove game-changing by improving genomic context in producing reads with better contiguity and more complete gene profiling [60], overcoming the challenge of assembling homopolymers. This could be key to uncovering functions encoded within large operons or ARG clusters [61] and more accurately detecting mutations, permitting detection of more novel ARGs [62]. Several long-read sequencing platforms are under development, two of which are currently widely used: single-molecule real-time sequencing (Pacific Biosciences) [62] and nanopore technology (Oxford Nanopore Technologies) [63]. Nanopore technology enables real-time base-calling, drastically reducing the sequencing time for a biological sample. However, both technologies are limited by throughput, cost, accuracy, and the processing limitations of current bioinformatic pipelines.

The utility of functional metagenomics

Functional metagenomics involves expression of a DNA sequence in a heterologous host to screen for a functional phenotype (Figure 2), bridging the divide between culture-based and genomic techniques, enabling identification of functional AMR from novel ARGs [23]. Functional metagenomics identifies only those ARGs with the capacity for HGT that are functional in a heterologous host – genes that are capable of transfer to, and expression in, pathogens [47]. Over the past decade, this technique has identified novel rifamycin phosphorylases [64] and tetracycline-inactivating enzymes [65]. Creation of functional metagenomic libraries, a process detailed elsewhere in the literature [47,66], could enable future proactive antibiotic development strategies by screening novel compounds for AMR [1,23,67].

The major limitations of functional metagenomics for the nasopharyngeal resistome are the low bacterial biomass of nasopharyngeal samples and the lack of phylogenetically diverse heterologous hosts [47,48]. Functional metagenomics can also miss ARGs arising from target mutations; if just one copy of the ARG with a target mutation is present, it can produce a dominant susceptible phenotype [1]. Functional metagenomics can also simulate HGT in otherwise immobile genetic elements (e.g., ARGs not collocated with an MGE) [1]. The limited insert size of a functional metagenomic library hinders its ability to identify combinatorial AMR mechanisms and MGEs, which have long genetic sequences [4]. Finally, comparison of functional screening techniques between laboratories is impeded by use of different media, incubation methods, and antibiotic concentrations [4]. Ultimately, this method has limited throughput, and is best combined with sequence-based techniques for resistome analysis [23]. In time, metatranscriptomics could replace current functional approaches [4].

Key factors in bioinformatic resistome analysis

The challenges of assembling metagenomic data

Metagenomic assembly is very challenging in low-biomass samples as the human genome is a thousand times larger than a bacterial genome, so can completely obscure low numbers of bacterial reads [19]. Metagenomic assembly is also complicated by uneven sequence coverage of multiple different organisms, with long stretches of identical sequences in unrelated species, making it difficult to assign reads to species. Algorithms attempt to account for this by using an optimised de Bruijn graph in which reads are divided into **k-mers** that form a network graph; assemblers reconstruct the genome sequence by finding the optimum (Euler's) path through this graph [59,68].

The plethora of software packages available for *de novo* assembly of metagenomic data is potentially overwhelming [59]. Despite attempts to benchmark these programs [69], there is no consensus on the optimal assembly package [59]. The current approach is to use multiple

Table 2. The challenges and potential solutions in metagenomic sequencing of bacterial samples for resistome analysis

Stage in sequencing process	Challenges	Ways to address these challenges	Refs
DNA extraction for resistome metagenomics	<ul style="list-style-type: none"> Ensuring equal lysis of Gram-positive and Gram-negative organisms for even species coverage within the resistome. Different DNA extraction kits with varying reagents, wash steps and incubation periods for enzymatic lysis can introduce bias. 	<ul style="list-style-type: none"> Choice of extraction technique: mechanical lysis improves overall yield, but shears DNA producing shorter reads that lack genomic context. Enzymatic lysis is not as effective for Gram-positive organisms, but produces longer reads. DNA extraction kits should be carefully selected for the intended sequencing technique, e.g., to optimise high-molecular-weight DNA for long-read sequencing. 	[46,70,82,83]
DNA extraction from low-biomass samples	<ul style="list-style-type: none"> Overwhelmed by human host DNA, which conceals bacterial signals, preventing detection of clinical pathogens and ARGs. Obtaining sufficient DNA volumes for sequencing. 	<ul style="list-style-type: none"> Depletion of human host DNA prior to or after sequencing improves bacterial DNA resolution, but both reduce overall DNA yield, and the latter is computationally intensive. Ultra-deep sequencing, which can be prohibitively expensive. Use of targeted amplification sequencing techniques e.g., multiplex PCR for 16S rDNA. 	[19,20,46,63,84]
Sequencing the bacterial resistome	<ul style="list-style-type: none"> High variation between samples requires multiple replications, which can be prohibitively expensive. Uneven species coverage in diverse resistome. 	<ul style="list-style-type: none"> No solution: remains a difficult balance between sufficient sequencing depth and number of samples that can be replicated. 'Binning' of coabundant species; sequences are classified into separate 'bins' (e.g., species) based on DNA composition patterns of the genomes (e.g., tetramer frequencies). This reduces assembly complexity and targets regions of interest. 	[46,76,85]

assembly tools to determine the best fit for a dataset; this is time-consuming and increases the already intensive computational demands of metagenomic assembly [59]. Metagenomic assembly typically requires hundreds of gigabytes of Random Access Memory, distributed across multiple processor cores and nodes in a cluster [46]. This requires considerable investment in supercomputing resources or cloud-based computing services with the data storage, management, and processing capabilities to cope with large volumes of data [4].

There are also specific challenges to assembling short reads for resistome analysis. HGT has led to sequence homology in unrelated species, and MGEs often contain homopolymers, long stretches of repeated sequences which are difficult to assemble [32,70]. Low-quality assembly can result in composite genomes that incorporate contigs from distinct populations, leading to flawed conclusions about microbial ecology [71]. This can be partially overcome by high-quality assembly with longer contigs, fewer gaps in the genome, and low error rates [32].

Sequence alignment and identity cut-offs

ARGs are identified from assembled contigs or genomes by comparison to reference databases using similarity based alignment search tools [59]. However, whilst closely related ARGs often contain identical regions, they will not be identical across their entire sequence length and may be split onto several shorter contigs, resulting in contigs containing partial sequences from resistance determinants [46].

Table 3. A comparison of currently available ARG databases

ARG databases	Updated/archived	Curated	Link	Description	Advantages	Disadvantages	Refs
Generalised databases – cover broad spectrums of ARGs and resistance mechanism information							
Antibiotic Resistance Gene Database (ARDB)	Archived	Yes	http://arcb.cbcb.umd.edu/	<ul style="list-style-type: none"> Uses BLAST to identify and annotate ARGs. 	<ul style="list-style-type: none"> The original ARG database, which was manually curated. 	<ul style="list-style-type: none"> Does not distinguish between putative ARGs based on homology and functionally confirmed ARGs. 	[46,59,86]
AMRFinderPlus	Updated	Yes	AMR Finder Plus: https://www.ncbi.nlm.nih.gov/pathogens/antimicrobial-resistance/AMRFinder/ NCBI's Reference Gene Database: https://www.ncbi.nlm.nih.gov/pathogens/refgene/# NCBI's Curated HMMs: https://ftp.ncbi.nlm.nih.gov/hmm/NCBIfam-AMRFinder/latest/	<ul style="list-style-type: none"> A tool hosted by NCBI, that uses the NCBI's Reference Gene Database and a curated collection of HMM. Repository of ARGs, resistance-associated point mutations, stress response and virulence genes. Contains an annotation tool applicable to assembled sequences, which can identify nucleotide sequence, contig location and internal stop codons. 	<ul style="list-style-type: none"> Permits inclusion/exclusion of putative stress response and virulence genes in search criteria. Permits detection of point mutations in protein and nucleotide sequences. Utilises manually curated BLAST out-ofs within a HMM function. 	<ul style="list-style-type: none"> Inclusion of putative ARGs, which lack functional validation. 	[87]
ResFinder	Updated	Yes	https://cge.cbs.dtu.dk/services/ResFinder/	<ul style="list-style-type: none"> Combines ARDB, Bush-Jacoby Beta-Lactamase list and other published ARG sequences. Uses BLAST pairwise sequence alignment tools for identification of acquired ARGs and point mutations based on the databases curated by Centre for Genomic Epidemiology (Denmark) 	<ul style="list-style-type: none"> Allows user to adjust the identify and length coverage thresholds. Applies a stringent approach to ARG identification; database only contains sequences of ARGs present on MGEs. 	<ul style="list-style-type: none"> Unable to classify chromosomal mutations causing antibiotic resistance or novel ARGs. Limited to ARGs acquired through HGT [46]. Does not distinguish between acquired and mutated ARGs [6]. 	[46,88–90]
SARG (version 2)	Updated		http://smile.hku.hk/SARGs	<ul style="list-style-type: none"> Hierarchical structured database derived from ARDB, CARD and NCBI-DR database. Used with ARGs-OAP (v2); a hybrid UBLAST and BLASTX algorithm. 	<ul style="list-style-type: none"> Comprehensive database with lower identity matching for ARG annotation of metagenomic sequence data. Comprehensive; contains >12 000 ARG and profile HMMs for 189 ARG subtypes. 	<ul style="list-style-type: none"> None identified. 	[59,91]
Comprehensive Antibiotic Resistance Database (CARD)	Updated	Yes	https://card.mcmaster.ca/	<ul style="list-style-type: none"> Accompanied by a BLAST-based pairwise sequence alignment tool. Organised by Antibiotic Resistance Ontology, a robust vocabulary and 	<ul style="list-style-type: none"> The most comprehensive resource for ARG available. Actively managed and periodically updated. Ideal database if 	<ul style="list-style-type: none"> Poor at detection of point mutations in chromosomal target genes known to be associated with ARGs. Curation based, but does 	[6,46,76,89,92]

(continued on next page)

Table 3. (continued)

ARG databases	Updated/archived	Curated	Link	Description	Advantages	Disadvantages	Refs
				framework for antibiotic resistance information. This permits different levels of organisation and relationships among sequences. Includes the Resistance Gene Identifier (RGI) a tool predicting ARGs from sequence data.	antibiotic resistance mechanism of interest is understood/well documented. <ul style="list-style-type: none"> • ARG information is linked to literature references, chemical and drug databases. • Uses a single reference sequence for each ARG, which increases likelihood of functional resistance confirmation in at least some species. 	not separate experimentally verified and predicted ARGs. <ul style="list-style-type: none"> • Does not differentiate between MGE-associated and chromosomal ARGs. 	
Functionally validated databases – only include functionally validated ARGs							
ResfinderFG	Updated; last November 2016	Yes	https://cge.cbs.dtu.dk/services/ResFinderFG/	<ul style="list-style-type: none"> • ARG variants identified in functional metagenomics studies. • Built by aggregating data from four functional metagenomics studies selected against 23 antibiotics. 	<ul style="list-style-type: none"> • Functionally validated ARGs. 	<ul style="list-style-type: none"> • Less comprehensive list of ARGs (only functionally validated), which may limit detection of novel resistance determinants. 	[59,93]
Resqu	Yes	Yes	https://www.1928diagnostics.com/resdb/	<ul style="list-style-type: none"> • Only contains ARGs with experimentally verified function and evidence of HGT between species. 	<ul style="list-style-type: none"> • More stringent approach to identifying ARGs; the database only contains sequences of acquired ARGs present on MGEs. • DNA sequences included in the database are manually assessed for errors, nonredundant and consistent with established gene families. 	<ul style="list-style-type: none"> • Consistently reports lowest numbers of ARGs. • Less comprehensive list of ARGs (only functionally validated), which may limit detection of novel resistance determinants. 	[46,94]
Resfamns	Updated; last in January 2015	Yes	http://www.dantaslab.org/resfams/	<ul style="list-style-type: none"> • Database of HMM profiles built on resistance proteins compiled from the CAPD, the LacED and Jacoby and Bush's collection of curated β-lactamases, based on the gene ontology from the CAPD database. • Curated database for ARGs with confirmed antibiotic resistance function. 	<ul style="list-style-type: none"> • HMMs identify novel ARGs with greater precision and accuracy than pairwise annotation. • These HMMs include a broader range of ARGs identified through culture-independent functional metagenomic selection. • HMM profiles are confirmed for functional resistance []. 	<ul style="list-style-type: none"> • Less comprehensive list of ARGs (only functionally validated), which may limit detection of novel resistance determinants. 	[6,23,59,74]

Functional Antibiotic Resistance Metagenomic Element (FARME) database	Updated; last in November 2017	Yes	http://staff.washington.edu/jwallace/farme/	<ul style="list-style-type: none"> Curated set of microbial sequences, excluded from current databases, but functionally screened to confer resistance in functional metagenomic studies. 	<ul style="list-style-type: none"> Includes protein--coding antibiotic resistance sequences, regulatory elements, MGE and predicted proteins flanking ARGs. 	<ul style="list-style-type: none"> Less comprehensive list of ARGs (only functionally validated), which may limit detection of novel resistance determinants. 	[59,95]
Mustard	Updated; last in November 2018	Yes	http://mgps.eu/Mustard/	<ul style="list-style-type: none"> Contains >6000 resistance determinants from 20 families. Includes a curated set of ARG identified in functional metagenomics studies. 	<ul style="list-style-type: none"> Functional validation of all ARGs included in the database. 	<ul style="list-style-type: none"> Less comprehensive list of ARGs (only functionally validated), which may limit detection of novel resistance determinants. 	[59,96]
Specialised databases – provide comprehensive information for specific gene families or species							
Beta-lactamase DataBase (BLDB)	Updated; last June 2022	Yes	http://bldb.eu/	<ul style="list-style-type: none"> Complex ARG sequence information, and biochemical and structural data for all currently known β-lactamases. 		<ul style="list-style-type: none"> Limited to a defined subset of ARGs. 	[59,97]
Lahey database of β -lactamases/-Bush-Jacoby Beta-Lactamase list	Archived	Yes	http://www.lahey.org/Studies/	<ul style="list-style-type: none"> Matches β-lactamase gene sequences with resistance to β-lactamase antibiotics. 	<ul style="list-style-type: none"> One of the first databases to compile known β-lactamases and assign nomenclature to new ones. 	<ul style="list-style-type: none"> Focussed on compiling a list of a small subset of ARGs. 	[59,76,98]
Lactamase Engineering Database (LacED)	TEMLacED active; last update in 2017; SHVED archived; last update in April 2010	Yes	http://www.laced.uni-stuttgart.de/	<ul style="list-style-type: none"> An extension of the Bush-Jacoby Beta Lactamase list. This database constructs protein families, using DWARF to integrate protein sequence and structure information. 	<ul style="list-style-type: none"> Includes methods to predict TEM, SHV, and class B enzymes by merging information from both the NCBI databases and TEM mutation table. Provides information on mutations, sequences and structures of TEM and SHV β-lactamases. 	<ul style="list-style-type: none"> Limited to a defined subset of ARGs. 	[6,76,99,100]
Repository of Antibiotic Resistance Cassettes	Yes	Yes	http://rac.ahh.mq.edu.au/rac/	<ul style="list-style-type: none"> Archive of ARG cassettes in multidrug--resistance' integrons. Uses proximity/synergy of MGEs with ARGs to form cassettes. 	<ul style="list-style-type: none"> First database to simultaneously automate identification of ARGs and place them in broader genetic context. 	<ul style="list-style-type: none"> Gene names differ from multiple other nomenclatures. Limited to a defined subset of ARGs. 	[6,101]

Basic Local Alignment Search Tool (BLAST) pairwise sequence alignment [72,73] is the most commonly used alignment approach [59] but is biased by the comparatively over-sampled human microbiome [74]. One alternative method is multiple sequence alignment using probabilistic prediction algorithms based on Hidden Markov Models (HMMs), enabling identification of a more diverse range of ARGs [23,74] with greater precision and accuracy than BLAST [75]. However, both BLAST and HMM statistics lack fine resolution, prompting difficult decisions regarding sequence identity cut-offs to prevent false-positive predictions [76]. Each technique can lead to differing and sometimes contradictory conclusions [74].

Cut-offs for alignment identity, score, and coverage breadth are important parameters to decide prior to ARG annotation, as they affect the number and type of ARGs detected [32]. Stringent sequence identity cut-offs avoid over-classification of ARGs, but the challenge is defining what is appropriate. Metagenomic reads will contain some sequencing errors, and not all genes with similar functions will have identical sequences. It is reasonable to allow some mismatch between read and reference sequences, but the degree of stringency in these mismatches affects the validity of results [46]. Accommodating mismatches when mapping short-read data also consumes computational resources [46]. Functional metagenomic studies help to validate identity cut-offs, but reads only represent part of a genetic sequence, so even if a read is 100% identical to an ARG it may not be functional [46]. The current solution for many resistome studies is a sequence identity cut off of 80–95%, a compromise between efficiency, sensitivity, and stringency of data analysis [46]. Some ARG databases have corresponding software packages for ARG annotation that provide gene-family specific alignment cut-offs [34].

ARG databases

There are multiple ARG reference databases (Table 3), compiled from studies of genetic determinants of AMR [59]. Databases that selectively include ARGs on MGEs identify resistance determinants most at risk of HGT, but miss novel and emerging ARGs [46]. Conversely, poorly curated databases that include chromosomal point mutations and putative ARGs, which lack functional validation, over-estimate the abundance and diversity of resistance determinants [6,46].

Some chromosomal point mutations in specific genes can produce AMR, for example, *rpoB*, but reads from wild-type genes will also map to these sequences due to their similarity. Many are ubiquitously occurring essential genes that dilute a database (e.g., PointFinder [77] and AMRFinder [78]), unless they are proven to confer functional resistance [32,46]. Therefore, a priority for the field is the functional validation of ARGs before database inclusion.

Another limitation of current ARG databases is the exclusion of environmental and nonclinical ARGs, limiting the ability to detect emerging resistance mechanisms in surveillance studies [74]. Many ARG databases also catalogue resistance determinants in a manner that ignores their taxonomic context, disregarding the host-specific functionality of some ARGs [76]; for example, ARGs for certain antibiotics in Gram-positive organisms may be irrelevant in Gram-negative organisms that are intrinsically resistant to those antibiotics, as they cannot penetrate their outer membrane. To address this, ARG databases will need to include species-specific risk estimations for functional resistance and HGT of ARGs [23].

Database curation

The challenge of ARG database curation impedes global data sharing and multinational surveillance of emerging AMR [63,76]. The best ARG databases are regularly updated with novel, functionally validated resistance determinants [1,23], but this requires manual curation and linking of complex metadata on resistance profiles and bacterial epidemiology [76]. A more automated method of database curation, which could operate through automatic incorporation of parameters from published data, is therefore urgently required if resistome analysis is to fulfil its potential.

Concluding remarks and future perspectives

AMR is placing an increasingly intolerable clinical and economic burden on global health systems, threatening a public health crisis on a scale that exceeds recent pandemics. Defining the nasopharyngeal resistome in human populations will improve AMR surveillance and enable the discovery of new ARGs and resistance mechanisms. It will also inform response strategies, including the repurposing of existing antibiotics into synergistic combinations, development of regimens that induce collateral sensitivity, and screening of novel compounds for resistance with functional metagenomic libraries.

The nasopharynx is an important, understudied, reservoir for emergence and spread of antibiotic-resistant bacteria that cause invasive disease. There are few studies of the nasopharyngeal resistome due to the challenges of characterising a resistome with a low sample biomass and high burden of human host DNA, alongside the general challenges of resistome characterisation, which include determining genetic context and phenotypic expression of ARGs, differing bioinformatic approaches and limited computational resources [63,79].

Advances in shotgun and long-read metagenomic sequencing techniques are providing tools to overcome these challenges, but require further development to improve reliability. ARG databases require optimisation to maximise their potential for AMR surveillance, including automated metadata curation and functional validation of all catalogued ARGs.

Current sequencing technologies could be used to characterise nasopharyngeal resistome evolution under specific selective pressures, for example antibiotic mass drug administration or change in pneumococcal vaccination regimen. However, for future nasopharyngeal resistome studies to fully inform preventative public health policy, for example the impact of mass drug administration and vaccination strategies on the nasopharyngeal resistome in children in high-disease-burden settings [80,81], there needs to be better linkage between ARGs and their bacterial host, and improved resolution of an individual's resistome from a single sample (see [Outstanding questions](#)).

Acknowledgments

L.O'C is supported by a British Infection Association (BIA) small project grant. R.S.H is supported by the National Institute for Health Research (NIHR) Global Health Research Unit on Mucosal Pathogens using UK aid from the UK Government. The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR, BIA or the Department of Health and Social Care.

Declaration of interests

No interests are declared.

References

- Crofts, T.S. *et al.* (2017) Next-generation approaches to understand and combat the antibiotic resistome. *Nat. Rev. Microbiol.* 15, 422–434
- O'Neill, J. (2016) Tackling drug-resistant infections globally: Final report and recommendations. *Arch. Pharm. Pract.* 7, 110
- Wright, G.D. (2010) The antibiotic resistome. *Expert Opin. Drug Discov.* 5, 779–788
- Schmieder, R. and Edwards, R. (2012) Insights into antibiotic resistance through metagenomic approaches. *Future Microbiol.* 7, 73–89
- Finley, R.L. *et al.* (2013) The scourge of antibiotic resistance: the important role of the environment. *Clin. Infect. Dis.* 57, 704–710
- Martinez, J.L. *et al.* (2015) What is a resistance gene? Ranking risk in resistomes. *Nat. Rev. Microbiol.* 13, 116–123
- Willmann, M. *et al.* (2019) Distinct impact of antibiotics on the gut microbiome and resistome: A longitudinal multicenter cohort study. *BMC Biol.* 17, 1–18
- Flynn, M. and Dooley, J. (2021) The microbiome of the nasopharynx. *J. Med. Microbiol.* 70, 001368
- Rutebemberwa, E. *et al.* (2015) High prevalence of antibiotic resistance in nasopharyngeal bacterial isolates from healthy children in rural Uganda: a cross-sectional study. *Ups. J. Med. Sci.* 120, 249–256
- Bojang, A. *et al.* (2020) Impact of intrapartum oral azithromycin on the acquired macrolide resistome of infants' nasopharynx: a randomized controlled trial. *Clin. Infect. Dis.* 71, 3222–3225
- Keenan, J.D. *et al.* (2018) The effect of antibiotic selection pressure on the nasopharyngeal macrolide resistome: a

Outstanding questions

Can we go beyond population-level analyses with current metagenomic technologies to draw conclusions about an individual's resistome and the potential for functional resistance in specific pathogens from a single sample?

Can long-read sequencing technologies fulfil their promise, enabling identification of ARGs on pathogenic species or colocalised with MGEs in the nasopharyngeal resistome?

Can data on the genomic context of ARGs identify which commensal bacteria are the major source of MGEs for pathogens in the nasopharyngeal resistome?

Can metagenomic sequencing be used to predict emerging AMR in environments with specific selective pressures, including:

- Clinical settings with high antibiotic use
- Community settings where mass drug administration has been used as a preventative public health policy
- Community settings where there has been a change in vaccine schedules that target pathogens within nasopharyngeal communities

- cluster-randomized trial. *Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am.* 67, 1736–1742
12. Pailhorès, H. *et al.* (2022) Antibiotic resistance in chronic respiratory diseases: from susceptibility testing to the resistome. *Eur. Respir. Rev.* 31, 210259
 13. Lagier, J.-C. *et al.* (2018) Culturing the human microbiota and culturomics. *Nat. Rev. Microbiol.* 16, 540–550
 14. Whelan, F.J. *et al.* (2020) Culture-enriched metagenomic sequencing enables in-depth profiling of the cystic fibrosis lung microbiota. *Nat. Microbiol.* 5, 379–390
 15. Li, L. *et al.* (2014) The importance of the viable but non-culturable state in human bacterial pathogens. *Front. Microbiol.* 5, 258
 16. Yang, C. *et al.* (2021) A review of computational tools for generating metagenome-assembled genomes from metagenomic sequencing data. *Comput. Struct. Biotechnol. J.* 19, 6301–6314
 17. Lanza, V.F. *et al.* (2018) In-depth resistome analysis by targeted metagenomics. *Microbiome* 6, 11
 18. Pehrsson, E.C. *et al.* (2016) Interconnected microbiomes and resistomes in low-income human habitats. *Nature* 533, 212–216
 19. Shi, Y. *et al.* (2022) Metagenomic sequencing for microbial DNA in human samples: emerging technological advances. *Int. J. Mol. Sci.* 23, 2181
 20. Nelson, M.T. *et al.* (2019) Human and extracellular DNA depletion for metagenomic analysis of complex clinical infection samples yields optimized viable microbiome profiles. *Cell Rep.* 26, 2227–2240.e5
 21. Jervis-Bardy, J. *et al.* (2015) Deriving accurate microbiota profiles from human samples with low bacterial content through post-sequencing processing of Illumina MiSeq data. *Microbiome* 3, 19
 22. Manenzhe, R.I. *et al.* (2020) Characterization of pneumococcal colonization dynamics and antimicrobial resistance using shotgun metagenomic sequencing in intensively sampled South African infants. *Front. Public Health* 8, 543898
 23. Adu-Oppong, B. *et al.* (2017) Genomic and functional techniques to mine the microbiome for novel antimicrobials and antimicrobial resistance genes. *Ann. N. Y. Acad. Sci.* 1388, 42–58
 24. Baquero, F. *et al.* (2009) Ecology and evolution of antibiotic resistance. *Environ. Microbiol. Rep.* 1, 469–476
 25. Stone, L.K. *et al.* (2016) Compounds that select against the tetracycline-resistance efflux pump. *Nat. Chem. Biol.* 12, 902–904
 26. Gonzales, P.R. *et al.* (2015) Synergistic, collaterally sensitive β -lactam combinations suppress resistance in MRSA. *Nat. Chem. Biol.* 11, 855–861
 27. Rosenkilde, C.E.H. *et al.* (2019) Collateral sensitivity constrains resistance evolution of the CTX-M-15 β -lactamase. *Nat. Commun.* 10, 618
 28. D'Costa, V.M. *et al.* (2011) Antibiotic resistance is ancient. *Nature* 477, 457–461
 29. Cox, G. and Wright, G.D. (2013) Intrinsic antibiotic resistance: Mechanisms, origins, challenges and solutions. *Int. J. Med. Microbiol.* 303, 287–292
 30. Blair, J.M.A. *et al.* (2015) Molecular mechanisms of antibiotic resistance. *Nat. Rev. Microbiol.* 13, 42–51
 31. Taylor, S.L. *et al.* (2018) Understanding the impact of antibiotic therapies on the respiratory tract resistome: a novel pooled-template metagenomic sequencing strategy. *Multidiscip. Respir. Med.* 13, 30
 32. Lee, K. *et al.* (2021) Overview of bioinformatic methods for analysis of antibiotic resistome from genome and metagenome data. *J. Microbiol.* 59, 270–280
 33. Nielsen, T.K. *et al.* (2022) Antibiotic resistance genes are differentially mobilized according to resistance mechanism. *Gigascience* 11, giac072
 34. Allemann, A. *et al.* (2019) Nasal resistome development in infants with cystic fibrosis in the first year of life. *Front. Microbiol.* 10, 212
 35. Taylor, S.L. *et al.* (2019) Long-term azithromycin reduces *Haemophilus influenzae* and increases antibiotic resistance in severe asthma. *Am. J. Respir. Crit. Care Med.* 200, 309–317
 36. Kalizang'oma, A. *et al.* (2021) *Streptococcus pneumoniae* serotypes that frequently colonise the human nasopharynx are common recipients of penicillin-binding protein gene fragments from *Streptococcus mitis*. *Microb. Genomics* 7
 37. Perry, J.A. *et al.* (2014) The antibiotic resistome: what's new? *Curr. Opin. Microbiol.* 21, 45–50
 38. Sultan, I. *et al.* (2018) Antibiotics, resistome and resistance mechanisms: A bacterial perspective. *Front. Microbiol.* 9, 2066
 39. Aminov, R.I. (2010) A brief history of the antibiotic era: lessons learned and challenges for the future. *Front. Microbiol.* 1, 134
 40. Cason, C. *et al.* (2021) Microbial contamination in hospital environment has the potential to colonize preterm newborns' nasal cavities. *Pathogens (Base)* 10, 615
 41. Manenzhe, R.I. *et al.* (2020) Longitudinal changes in the nasopharyngeal resistome of South African infants using shotgun metagenomic sequencing. *PLoS One* 15, e0231887
 42. Garcia-Rodríguez, J.A. and Fresnadillo Martínez, M.J. (2002) Dynamics of nasopharyngeal colonization by potential respiratory pathogens. *J. Antimicrob. Chemother.* 50, 59–73
 43. Unger, H.W. *et al.* (2015) Impact of intermittent preventive treatment in pregnancy with azithromycin-containing regimens on maternal nasopharyngeal carriage and antibiotic sensitivity of *Streptococcus pneumoniae*, *Haemophilus influenzae*, and *Staphylococcus aureus*: a cross-sectional study. *J. Clin. Microbiol.* 53, 1317–1323
 44. Hendriksen, R.S. *et al.* (2019) Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nat. Commun.* 10, 1124
 45. Hays, J. (2011) Mobile genetic elements in *Moraxella catarrhalis*. *Mob. Genet. Elem.* 1, 155–158
 46. Bengtsson-Palme, J. *et al.* (2017) Using metagenomics to investigate human and environmental resistomes. *J. Antimicrob. Chemother.* 72, 2690–2703
 47. Lam, K.N. *et al.* (2015) Current and future resources for functional metagenomics. *Front. Microbiol.* 6, 1196
 48. Pehrsson, E.C. *et al.* (2013) Novel resistance functions uncovered using functional metagenomic investigations of resistance reservoirs. *Front. Microbiol.* 4, 145
 49. Cox, M.J. *et al.* (2017) Longitudinal assessment of sputum microbiome by sequencing of the 16S rRNA gene in non-cystic fibrosis bronchiectasis patients. *PLoS One* 12, e0170622
 50. Mac Aogáin, M. *et al.* (2020) Metagenomics reveals a core macrolide resistome related to microbiota in chronic respiratory disease. *Am. J. Respir. Crit. Care Med.* 202, 433–447
 51. Bojang, A. *et al.* (2018) Long-term impact of oral azithromycin taken by Gambian women during labor on prevalence and antibiotic susceptibility of *Streptococcus pneumoniae* and *Staphylococcus aureus* in Their Infants: follow-up of a randomized clinical trial. *Clin. Infect. Dis.* 67, 1191–1197
 52. Haug, S. *et al.* (2010) The decline of pneumococcal resistance after cessation of mass antibiotic distributions for trachoma. *Clin. Infect. Dis.* 51, 571–574
 53. Roca, A. *et al.* (2016) Oral azithromycin given during labour decreases bacterial carriage in the mothers and their offspring: a double-blind randomized trial. *Clin. Microbiol. Infect.* 22, 565.e1–565.e9
 54. Bojang, A. *et al.* (2019) Genomic investigation of *Staphylococcus aureus* recovered from Gambian women and newborns following an oral dose of intra-partum azithromycin. *J. Antimicrob. Chemother.* 74, 3170–3178
 55. Yameen, M.A. *et al.* (2013) Nasal and perirectal colonization of vancomycin sensitive and resistant enterococci in patients of paediatrics ICU (PICU) of tertiary health care facilities. *BMC Infect. Dis.* 13, 156
 56. Rajar, P. *et al.* (2022) Microbial DNA extraction of high-host content and low biomass samples: optimized protocol for nasopharynx metagenomic studies. *Front. Microbiol.* 13, 1038120
 57. Mamanova, L. *et al.* (2010) Target-enrichment strategies for next-generation sequencing. *Nat. Methods* 7, 111–118
 58. Raghavendra, P. and Pullaiah, T. (2018) Pathogen Identification Using Novel Sequencing Methods. In *Advances in Cell and Molecular Diagnostics* (Raghavendra, P. and Pullaiah, T., eds), pp. 161–202, Elsevier
 59. Boolchandani, M. *et al.* (2019) Sequencing-based methods and resources to study antimicrobial resistance. *Nat. Rev. Genet.* 20, 356–370
 60. Metzker, M.L. (2010) Sequencing technologies – the next generation. *Nat. Rev. Genet.* 11, 31–46
 61. Koren, S. *et al.* (2013) Reducing assembly complexity of microbial genomes with single-molecule sequencing. *Genome Biol.* 14, R101

62. Rhoads, A. and Au, K.F. (2015) PacBio sequencing and its applications. *Genomics Proteomics Bioinformatics* 13, 278–289
63. Didelot, X. *et al.* (2012) Transforming clinical microbiology with bacterial genome sequencing. *Nat. Rev. Genet.* 13, 601–612
64. Spanogiannopoulos, P. *et al.* (2014) A rifamycin inactivating phosphotransferase family shared by environmental and pathogenic bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 111, 7102–7107
65. Forsberg, K.J. *et al.* (2015) The tetracycline destructases: a novel family of tetracycline-inactivating enzymes. *Chem. Biol.* 22, 888–897
66. Boolchandani, M. *et al.* (2017) Functional metagenomics to study antibiotic resistance. *Methods Mol. Biol.* 1520, 307–329
67. Baym, M. *et al.* (2016) Multidrug evolutionary strategies to reverse antibiotic resistance. *Science* 351, 6268
68. Lapidus, A.L. and Korobeynikov, A.I. (2021) Metagenomic data assembly – the way of decoding unknown microorganisms. *Front. Microbiol.* 12, 613791
69. Meyer, F. *et al.* (2022) Critical assessment of metagenome interpretation: the second round of challenges. *Nat. Methods* 19, 429–440
70. Laudadio, I. *et al.* (2019) Next-generation metagenomics: methodological challenges and opportunities. *OMICS* 23, 327–333
71. Chen, L.X. *et al.* (2020) Accurate and complete genomes from metagenomes. *Genome Res.* 30, 315–333
72. Altschul, S.F. *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402
73. Altschul, S.F. *et al.* (1990) Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410
74. Gibson, M.K. *et al.* (2015) Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. *ISME J.* 9, 207–216
75. Pesesky, M.W. *et al.* (2016) Evaluation of machine learning and rules-based approaches for predicting antimicrobial resistance profiles in gram-negative bacilli from whole genome sequence data. *Front. Microbiol.* 7, 1887
76. McArthur, A.G. and Wright, G.D. (2015) Bioinformatics of antimicrobial resistance in the age of molecular epidemiology. *Curr. Opin. Microbiol.* 27, 45–50
77. Zankari, E. *et al.* (2017) PointFinder: a novel web tool for WGS-based detection of antimicrobial resistance associated with chromosomal point mutations in bacterial pathogens. *J. Antimicrob. Chemother.* 72, 2764–2768
78. Feldgarden, M. *et al.* (2019) Validating the AMRFinder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. *Antimicrob. Agents Chemother.* 63 e00483-19
79. Wright, M.S. *et al.* (2017) Strain Level Streptococcus Colonization Patterns during the First Year of Life. *Front. Microbiol.* 8, 1661
80. Keenan, J.D. *et al.* (2018) Azithromycin to reduce childhood mortality in sub-Saharan Africa. *N. Engl. J. Med.* 378, 1583–1592
81. Swarouth, T.D. *et al.* (2020) High residual carriage of vaccine-serotype *Streptococcus pneumoniae* after introduction of pneumococcal conjugate vaccine in Malawi. *Nat. Commun.* 11, 1–12
82. Trigodet, F. *et al.* (2022) High molecular weight DNA extraction strategies for long-read sequencing of complex metagenomes. *Mol. Ecol. Resour.* <https://doi.org/10.1111/1755-0998.13588>
83. Greathouse, K.L. *et al.* (2019) DNA extraction for human microbiome studies: the issue of standardization. *Genome Biol.* 20, 212
84. Zhang, C. *et al.* (2019) Use of ultra-deep sequencing in a patient with tuberculous coxitis shows its limitations in extrapulmonary tuberculosis diagnostics: a case report. *Infect. Drug Resist.* 12, 3739–3743
85. Plaza Oñate, F. *et al.* (2019) MSPminer: abundance-based reconstruction of microbial pan-genomes from shotgun metagenomic data. *Bioinformatics* 35, 1544–1552
86. Liu, B. and Pop, M. (2009) ARDB – antibiotic resistance genes database. *Nucleic Acids Res.* 37, D443–D447
87. Feldgarden, M. *et al.* (2021) AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. *Sci. Rep.* 11, 12728
88. Zankari, E. *et al.* (2012) Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* 67, 2640–2644
89. Gupta, S.K. *et al.* (2014) ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob. Agents Chemother.* 58, 212–220
90. Bortolaia, V. *et al.* (2020) ResFinder 4.0 for predictions of phenotypes from genotypes. *J. Antimicrob. Chemother.* 75, 3491–3500
91. Yin, X. *et al.* (2018) ARGs-OAP v2.0 with an expanded SARG database and Hidden Markov Models for enhancement characterization and quantification of antibiotic resistance genes in environmental metagenomes. *Bioinformatics* 34, 2263–2270
92. McArthur, A.G. *et al.* (2013) The comprehensive antibiotic resistance database. *Antimicrob. Agents Chemother.* 57, 3348–3357
93. Munk, P. *et al.* (2018) Abundance and diversity of the faecal resistome in slaughter pigs and broilers in nine European countries. *Nat. Microbiol.* 3, 898–908
94. 1928 Diagnostics Resqu: a database of mobile antibiotic resistance genes. [Online]. Available: <https://www.1928diagnostics.com/resdb/>, Accessed date: 21 March 2022
95. Wallace, J.C. *et al.* (2017) FARME DB: a functional antibiotic resistance element database. *Database* 2017, baw165
96. Ruppé, E. *et al.* (2019) Prediction of the intestinal resistome by a three-dimensional structure-based method. *Nat. Microbiol.* 4, 112–123
97. Naas, T. *et al.* (2017) Beta-lactamase database (BLDB) - structure and function. *J. Enzyme Inhib. Med. Chem.* 32, 917–919
98. Bush, K. and Jacoby, G.A. (2010) Updated functional classification of beta-lactamases. *Antimicrob. Agents Chemother.* 54, 969–976
99. Fischer, M. *et al.* (2006) DWARF – a data warehouse system for analyzing protein families. *BMC Bioinformatics* 7, 495
100. Thai, Q.K. *et al.* (2009) The Lactamase Engineering Database: a critical survey of TEM sequences in public databases. *BMC Genomics* 10, 390
101. Tsafnat, G. *et al.* (2011) RAC: repository of antibiotic resistance Cassettes. *Database (Oxford)* 2011, bar054