COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

Review

Antibiotic resistance: Time of synthesis in a post-genomic age

Teresa Gil-Gil [a], Luz Edith Ochoa-Sánchez [a], Fernando Baquero [b,c], José Luis Martínez [a,*]

[a] Centro Nacional de Biotecnología, CSIC, Darwin 3, 28049 Madrid, Spain
[b] Department of Microbiology, Hospital Universitario Ramón y Cajal (IRYCIS), Madrid, Spain
[c] CIBER en Epidemiología y Salud Pública (CIBER-ESP), Madrid, Spain

ARTICLE INFO

ABSTRACT

Antibiotic resistance has been highlighted by international organizations, including World Health Organization, World Bank and United Nations, as one of the most relevant global health problems. Classical approaches to study this problem have focused in infected humans, mainly at hospitals. Nevertheless, antibiotic resistance can expand through different ecosystems and geographical allocations, hence constituting a One-Health, Global-Health problem, requiring specific integrative analytic tools. Antibiotic resistance evolution and transmission are multilayer, hierarchically organized processes with several elements (from genes to the whole microbiome) involved. However, their study has been tradition-ally gene-centric, each element independently studied. The development of robust-economically affordable whole genome sequencing approaches, as well as other -omic techniques as transcriptomics and pro-teomics, is changing this panorama. These technologies allow the description of a system, either a cell or a microbiome as a whole, overcoming the problems associated with gene-centric approaches. We are cur-rently at the time of combining the information derived from -omic studies to have a more holistic view of the evolution and spread of antibiotic resistance. This synthesis process requires the accurate integration of -omic information into computational models that serve to analyse the causes and the consequences of acquiring AR, fed by curated databases capable of identifying the elements involved in the acquisition of resistance. In this review, we analyse the capacities and drawbacks of the tools that are currently in use for the global analysis of AR, aiming to identify the more useful targets for effective corrective interventions.
© 2021 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creative-commons.org/licenses/by-nc-nd/4.0/).

Contents

* Corresponding author.
  E-mail address: jlmtnez@cnb.csic.es (J.L. Martínez).

1. Introduction

Infections had been historically the most important cause of human death. In this regard, the discovery and further use of antibiotics, and later on antifungal, antiviral and antiparasitic compounds, has been one of the major pharmacotherapeutic contributions to human welfare. The increasing prevalence of antibiotic resistant organisms is hence a relevant problem that has been recognized by several international agencies including, among others, the World Health Organization [1–4], United Nations (https://www.un.org/pga/71/wp-content/uploads/sites/40/2016/09/DGACM_GAEAD_ESCAB-AMR-Draft-Political-Declaration-1616108E.pdf) and the World Bank (https://www.worldbank.org/en/topic/health/publication/drug-resistant-infections-a-threat-to-our-economic-future). Studies on antibiotic resistance (AR) had traditionally focused on hospitals and health care centres. Further, they had been mainly based on phenotypic tests directed to identify the most appropriate antibiotics for treating the infection suffered by a given patient. Actually, AR is defined at clinical setting using breakpoints of minimal inhibitory concentrations (MICs), which are established based on the likelihood of the success of the treatment [5], no matter the resistance mechanisms involved. Despite its fundamental utility for treating individual patients, it has been revealed that this type of phenotypic tests is not potent enough for advanced epidemiological studies, including the analysis of outbreaks at hospitals. Molecular-based approaches as Multi Locus Sequence Typing (MLST) [6] or typing based on Pulse Field Gel Electrophoresis [7] demonstrated to be useful in the molecular epidemiology analysis or organism causing infections, while the implementation of PCR techniques was a breakthrough for tracking the dissemination of resistance genes and the mobile elements involved in their dispersal [8–11]. However, each of these techniques allowed to study just a limited aspect of the AR problem. In this regard, it is worth mentioning that AR is an evolutionary process in which different hierarchically interconnected elements participate [12,13]. An AR gene (ARG) can be acquired by an integron that is further included into a plasmid present in a given bacterium, that can transfer the plasmid to other organisms, each one capable of infecting a different host, which can also transmit the resistant organism to other hosts (Fig. 1) [14,15]. Analysing each of these levels of complexity and integrating the information to get a clear picture of the roads for the emergence

and transmission of AR is not an easy task [16,17]. Nevertheless, holistic tools, capable of analysing all complexity levels ideally without the need of isolating the resistant organisms, are making this difficult task more feasible.

Once the first genome was sequenced, *Haemophilus influenzae* in 1995, the advances in new technologies, both experimental and computational, for studying biological systems have been enormous [18], ending up in the development of systems biology approaches [19]. Instead of dividing a biological process into pieces as classical molecular biology approaches do, the aim of systems biology is to understand the functioning of a biological system as a whole. In the specific case of AR, the "selfish gene" concept of Dawkins [20] is generally applied; gene-centric (or mutation-centric) approaches are the current rule in the field. However, it is important to recall that the introduction of a new ARG or the selection of AR mutations can produce deep effects in bacterial physiology, altering sometimes the expression of hundreds of genes [21–23] and hence modifying the bacterial proteome and its metabolism. Besides, it is also relevant to state that the effect of one ARG (or AR mutation) on AR and on bacterial physiology is context-dependent; different in different organisms [24,25]. Therefore, in order to have a complete understanding of AR mechanisms and the consequences of acquiring resistance for bacterial physiology, bioinformatics tools and techniques able to analyse the increasing data generated in molecular biology through multi-omics technologies - including transcriptomics, proteomics and metabolomics - have become essential. The main objective of the information acquired through "-omics" is to create knowledge from data, providing simultaneous information on the presence and identification of alleles of thousands of genes, analysis of their genetic linkage with other genes, and measurements of their levels of expression as well as of the proteins they encode and cell metabolites [26]. To get this goal, system-scale models able to collect the huge amount of experimental information into mathematical models have been developed to further understand the bacterial cell as a whole [27,28].

Besides their contribution for analysing the causes and consequences of the acquisition of AR [29], the emergence of easy and non-expensive whole-genome sequencing (WGS) tools (also including transcriptomics studies) had been a hallmark for the study of the phylogenomics of bacterial pathogens [30,31] and AR [32–34]. These approaches are currently fundamental in
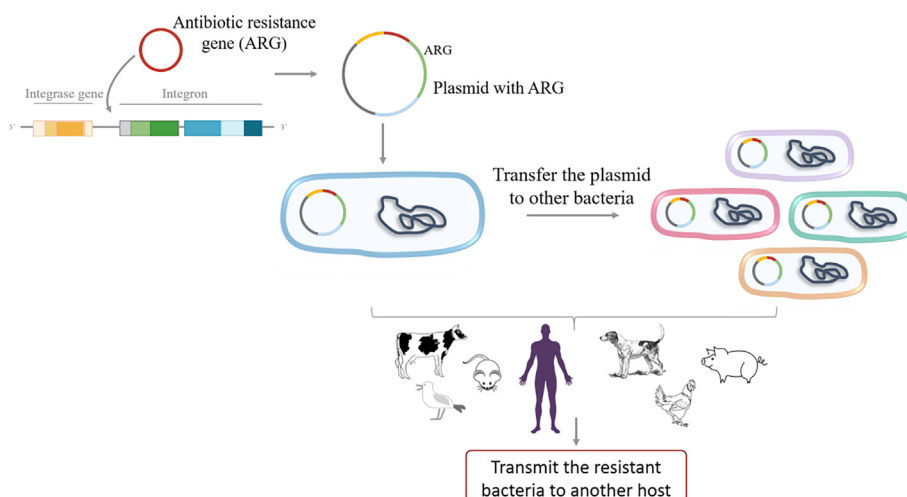
Fig. 1. AR Transmission as a multi-layered and hierarchical evolutionary process. ARGs are recruited by gene-capture elements such as integrons that are included in mobile elements as plasmids, which may be in turn acquired by a given bacterium. These clones can transfer the plasmid to other bacteria through HGT; these resistant bacteria may infect different hosts and then can spread among different hosts and environments. Thus, the process of transmission facilitates the evolution of AR at multiple, interconnected levels.

advanced molecular epidemiology studies (including the analysis of human-linked and environmental microbiomes) and had been also proposed as useful diagnostic tools. In the present review, we discuss current information on the advances and limitations of the use of these techniques for studying AR, under the umbrella of One-Health, Global-Health approaches [14,35,36].

2. Emergence and spread of antibiotic resistance in a multilevel selection space

By definition, many human bacterial pathogens were susceptible to the antibiotics used in therapy when these antibiotics started to be used (otherwise, the antibiotics would had not been developed by pharma companies). However, resistant variants began to emerge soon after antibiotics were introduced in clinical practice, either due to mutations [37] or as a consequence of the acquisition of ARGs [38]. With few exceptions [39–41], AR mutations are transferred only vertically; they are relevant just for the clones harbouring the mutation. In this case, clonal selection, and the associated clonal expansion, is the main element involved in AR spread. One aspect to address here is knowing if the acquisition of resistance favours specifically the enrichment of AR clones over the susceptible ones, hence altering the bacterial population structure, or if the most abundant clones are those more prone to acquire AR just by chance, in which case resistance will not necessarily alter the structure of the population. The situation in the case of ARGs acquired by Horizontal Gene Transfer (HGT) is different than in the case of mutational events since, as above stated, their selection implies multiple layers of organization (Fig. 1). Each layer is a unit of selection and, as a matter of fact, the properties of reproduction, inheritance, variation and interaction actually define an individual [42]. Each one of these individuals are the subject of Darwinian evolution. However, since they are hierarchically linked, selection-driven changes [43] in one of them will influence top-down (from microbiomes to genes) or bottom-up (from ARGs to clonal complexes) any one of the others, forming what has been dubbed as a nested living system [44]. Eventually, abundant clones that can acquire ARGs more easily might contribute to the homeostasis of the clonal structure of the species by giving access to minority clones to this "common ARG pool" by means of HGT.

Mutation driven resistance is a problem for the patient infected by bacteria presenting such mutations. Hence, it can become a global public health problem as a consequence of the clonal expansion of such mutants. While this is an obviously dangerous situation, the acquisition, through HGT, of AR determinants can be even more problematic. In addition to clonal expansion, the presence of ARGs in transferable elements allows their spread among different clones and even among different (frequently phylogenetically related) bacterial species [45–47]. One important aspect of the field of AR applied to human health is risk analysis. The risk of finding a given ARG in non-pathogenic bacteria is not as high as if the same gene is found in a virulent organism [5,48]. While the detection of gene-host association is easy in the case of isolated bacteria, the situation is not the same when microbiomes are analysed (see below).

Nevertheless, even when isolated bacteria are studied, the impact on human health of AR might be different according to the clone harbouring the ARG and the location of this gene (non-mobile or within a mobile element). Different bacterial isolates from the same species may have different niches and different properties. One example of this situation is Escherichia coli that, despite being a regular human commensal, presents some virulent clones, due to the incorporation in their genome of virulence determinants acquired by HGT. Further, the host range of pathogenic variants can vary; there are clonal complexes able to colonize/in-

fect humans and other animals and there are clones that are animal-specific [49,50]. When the problem for human health of selection of AR at farms is studied, this feature and the identification of shuttle clones or mobile elements capable of moving from animals to humans has to be taken into consideration [14]. It is important to highlight that the number of genes shared by all members of a given species (the core genome) is much lower than the total number of genes (the pangenome) that the species presents as a whole [51,52]. Knowing that different clones may behave differently concerning human health, different methodologies have been applied to identify those that most frequently cause infections and/or are involved in the dissemination of ARGs [53,54]. Currently, the most popular method for phylogenetic analysis is MLST [55], which is based on the identification of specific alleles with phylogenetic value in a relatively low number (typically 7–8) of housekeeping, highly conserved, genes [6]. MLST allows to define sequence types (ST), which are currently used to compare the phylogeny of isolates from different bacterial species all around the word [56,57]. Comparability of the results among different laboratories is based on the standardization of the methods and loci used for defining STs and in the existence of broad databases as PubMLST (https://pubmlst.org), which contains hundreds of thousands of genomes and millions of alleles. Despite its utility for inferring core-genome based phylogenies, the contribution of the accessory genome, which is critical for the bacterial adaptation to different habitats, cannot be assessed with this technology. In fact, while genome trees based on core genome are consistent with the species ontology, the phylogenetic relationships derived from the analysis of all genes (both core and accessory ones) are better represented as networks [58,59]. This does not mean that classical MLST analysis is lacking usefulness in a whole WGS era. When linked to PCR-based detection of ARGs, as well as to plasmid typing [8,60], these studies allow the multilevel analysis of the epidemiology of AR. Note that AR tends to spread and evolve in closely related bacteria lineages, eventually composed by assemblies of particular clones within [61]. In pathogenic bacteria, these clonal complexes frequently constitute the "high-risk" populations for human and animal health, and can be identified by MLST procedures. These types of analyses, mainly based on PCR approaches and Sanger-sequencing of the amplicons, still remain as the golden standard for analysing outbreaks and in regular molecular epidemiology studies dealing with the dissemination of AR. However, as will be discussed later on, WGS price reduction and the development of user-friendly bioinformatic tools for analysing sequencing results in a short time-lapse is changing this panorama.

3. Genomic approaches for the study of antibiotic resistance

The appearance of genomics, proteomics and transcriptomics has revolutionized the study of microorganisms. One of the fields that has benefited the most is the study of bacterial pathogens and their resistance to antimicrobials. Nowadays, the sequencing of a bacterial genome (or a metagenome) is affordable and the main constrains for applying this approach to the clinical field are the speed of sequencing (most laboratories outsource WGS) and the time and the skills required for the analysis of the obtained sequence. Although several of the analytical tools require some expertise in informatics, user-friendly bioinformatic approaches are increasingly implemented, which will favour these studies. However, while they are already useful for fundamental studies on AR and for epidemiological purposes [62], WGS-based approaches still have some limitations (see below) regarding their application in fast diagnostic procedures.

The main benefit of WGS-based approaches in the study of AR is that they can address simultaneously all the elements involved in

Table 1
Tools for analysing antibiotic resistance from genomes and metagenomes.

Tool	Tool type	Database-link	Access	Approach	Status	Input description	Requirements	References
AMRFinderPlus	Detection-database-based	Reference Gene Catalog	Web and standalone	NA; EA; Assembly-based and read based tool	Active	Protein search, protein.fa and nucleotide	HMMER, BLAST+, Linux, and Perl	[78]
ResFinder	Detection-database-based	resfinder_db	Web and standalone	NA; EA; Assembly-based and read based tool	Active	Whole genome sequencing, isolate or annotated genome, preassembled partial, complete genomes, reads		[79]
PointFinder	Detection-database-based	Pointfinder_db	Web and standalone	NA; EA; Assembly-based and read based tool	Active	Sequence file in FASTA	BioEdit platform (http://www.mbio.ncsu.edu/bioedit/)	[80]
Antibiotic Resistance Gene-ANNOTation (ARG-ANNOT)	Detection-database-based	ARG-ANNOT	Standalone	NA; EA; Assembly-based and read based tool	Active	Analysing genomes, genomes assemblies, metagenomic contigs or proteomes	Prodigal, DIAMON	[81]
The Resistance Gene Identifier (RGI)	Detection-database-based	CARD	Web and standalone	NA; EA; Assembly-based and read based tool	Active	Metagenomic	R	[82]
Online Analysis Pipeline for Anti-biotic Resistance Genes Detection (ARGs-OAP v2)	Detection-database-based		Web and standalone	NA; EA; Assembly-based and read based tool	Active	Sequence reads, allele db, gene db	python (v2.7.5)	[83]
Short Read Sequence Typing for Bacterial Pathogens (SRST2)	Detection-database-based		Standalone	NA; EA Read-based tool	Active	Metagenomic and db	Unix server with ~2 GB of disk space for reference data 2X the input FASTQ file size in both RAM and disk space for temporary file storage : Perl , r, usearch, bwa, tophat incl., bam2fastx, samtools incl. bcftools and vcfutils.pl, ncbi blast, seqtk	[84]
Search Engine for Antimicrobial Resistance (SEAR)	Detection-database-based		Web and standalone	NA; EA Read-based tool	Last update 2017	Sequence reads	Linux	[85]
Antimicrobial Resistance Identification By Assembly (ARIBA)	Detection-database-based	ARG-ANNOT, CARD, MEGAres and ResFinder	Standalone	NA; EA; Assembly-based and read based tool	Active	Sequence reads	Linux	[86]
SSTAR	Detection-database-based	SEED (annotator), AMR-related proteins that have been curated at ARDB and CARD	Standalone	NA; EA assembly-based and read based tool	Active	Two sequence files in FASTA format, one containing the bacterial genome assembly and the other the AR gene collection	NS	[73]
AdaBoost (PATRIC)	Detection and classification		Web	ML; EA Read-based tool	Active	Whole genome sequencing reads	NA	[87]
PARGT	Detection-database-based		Standalone	NA	Active	Protein sequences	R and Python	[88]
Resfams	Databased and AMR protein predictor	Resfams	Web and standalone	NA	Last update 2018	NA	None	[89]
Antibiotic Resistance Genes Database (ARDB)	Database	ARDB	Web and standalone	NA	Available	NA	None	[90]
NCBI Bacterial Antimicrobial Resistance	Database	PRJNA313047	Web	NA	Active	NA	None	

Table 1 (*continued*)

Tool	Tool type	Database-link	Access	Approach	Status	Input description	Requirements	References
Reference Gene Database (BARRGD)								
NCBI Pathogen Detection	Database	Pathogens	Web	NA	Active	NA	None	
National Database of Antibiotic Resistant Organisms (NDARO)	Database	NDARO	Web	NA	Active	NA	None	
Isolates Browser	Database	Insolates Browser	Web	NA	Active	NA	None	
RefSeq	Database	refseq	Web	NA	Active	NA	None	[91]
How to Request New Alleles for Beta-Lactamase, MCR, and Qnr Gene	Database	Home page	Web	NA	Active	NA	None	
Microbial Browser for Identification of Genetic and Genomic Elements (MicroBIG-E)	Database	microbigge	Web	NA	Active	NA	None	
HMDARG	Database	HMDARG	Web	Assembly-based tool and hierarchical classification	Active	Raw sequence encoding	None	[120]

NA: Not applicable EA: Exploratory approaches.

the acquisition of the phenotype and can establish the linkage of each ARG with the element harbouring it, either mobile or non-mobile. Further, WGS allows to develop genome-wide studies based on the analysis of several genes/mutations [63–70]. Furthermore, WGS is fundamental for performing phylogenomic studies including the accessory genome that go beyond classical MLST analyses, which as stated above, take into consideration just the core genome. Since bacterial evolution, and specifically the evolution of virulence and of AR, largely relies on the acquisition of virulence determinants/ARGs [71,72], the availability of WGS tools is particularly relevant for understanding the elements driving the evolution of virulence and AR. Besides the aspects dealing with epidemiology, which will be discussed later on, WGS approaches are useful tools not only to describe novel mechanisms of resistance but also to predict how resistance (even to drugs not present yet in the market yet) can emerge. To note here, epidemiological studies must necessarily rely on what is known. The identification of novel resistance mechanisms is frequently cumbersome and WGS approaches are helping to solve this problem. Indeed, these technologies are increasingly being used for the identification of genes associated with AR [73]; allowing the development of genome collections and the annotation of genes from clinically relevant strains [74]. Bioinformatics approaches to study antimicrobial resistance include computing programs able to predict, detect, infer and analyse ARGs from isolated culture data. These programs are used jointly with the systematic compilations of previous knowledge taken from databases (Table 1). Unlike other prediction methods, new detection methods employ rule-based inference, which takes advantage of already known AR phenotypes features. There are other methods that combine knowledge-based detection with machine-learning and mathematical inference approaches (hybrid methods) [75]. Specifically, Resfams addresses this remote homology detection problem employing hidden Markov models for AR protein identification [76]. Despite the potency of these methods, it is important to point out that predictions require validations.

When current methods for predicting AR are compared [76], they differ in aspects as their link to previous knowledge sources (database-based methods), tool access (standalone or web service) and the range of data used for the analysis (from gene or protein sequences to whole genome or proteome data). Several database-based prediction methods are assembly-based; however, some applications admit raw sequences or protein sequences as input data. While for analysing the structure of the genetic elements harbouring ARGs contig generation is needed, the use of raw sequences can be more appropriate for quantification purposes. Standardization of methodologies and databases is still ongoing, although regulatory agencies as FDA have raised initiatives in this direction [77].

As stated in [33] and credited to Gerry Wright from McMaster University, ARGs can be classified in known knowns (those already analysed), known unknowns (those that have not been analysed yet, but are homologous to the knowns) and unknown unknowns. Although AR can be context-dependent [24], we have the tools needed to link genotype with phenotype for the first category, at least in most cases. Reads-based methods use different strategies for detecting ARGs from raw reads. For example, SRST2 [84] and GeneFinder [92] use efficient read alignment; and Mykrobe [93] creates Bruijn graphs of contigs from raw data and matches them against known ARGs; however, it omits the generation of a consensus sequence. Regarding the second category of known unknowns, machine-learning approaches enable strong predictions since these algorithms work by finding the relevant features in complex datasets. This approach uses a "rules-based" classification based on the presence of one or more known ARGs or mutations [76].

The PATRIC annotation system links bacterial genomes with AR metadata, which favours genotype/phenotype linkage. This project has also implemented a prediction method called AdaBoost, a machine learning classifier that identifies specific ARGs using whole-genome sequences. A balanced dataset between susceptible and resistant genomes influences the algorithm accuracy. Unfortunately, there are insufficient numbers of genomes from clearly susceptible bacteria to build these classifiers because several sequenced organisms are resistant to antibiotics, due to their epidemiological relevance [94].

PARGT is another machine learning prediction method. It uses protein characteristics to identify ARGs in Gram-negative and Gram-positive bacteria. Results obtained from a simulation carried out by the authors showed that PARGT could predict AMR sequences in Gram-positive bacteria with 87 to 90% accuracy. Since

not all AR determinants contribute equally to the final AR phenotype, noise in phenotype predictions can often be reduced and accuracy increased by weighting each locus using a machine-learning model. These models can also be trained to consider potential interactions between loci. It is worth noting that there is compensation between the accuracy and the computational time when large-scale reference databases are used to generate the models [88]. Despite these advances in predictive methods, we must be aware that a single mutation can change the activity profile of an ARG, as it happens in the case of extended spectrum beta-lactamases [95]. Besides, synonymous mutations might have cryptic effects on the evolution of a global phenotype (resistance, fitness cost, and compensatory evolution), hence forming part of the "whisper mutations" set [96,97]. Altogether, this implies that a phenotype cannot be always directly inferred from a consensus sequence. However, once the novel mutational variant is detected, analysing the associated phenotype, and incorporate this information in the analysis of AR, is an easier task. The main problem lies in the unknown unknowns category and, despite advances in the field, novel ARGs that may emerge always require experimental validation of their activity. Examples of this situation are *qnr* [98] or *mcr-1* [99] conferring resistance to quinolones and polymyxin respectively and presenting sequences that did not resemble, at the time of their discovery, any previously known ARG.

Before WGS approaches were available, the identification of AR mutations relied on mechanisms-knowledge-driven approaches. Mutations in antibiotic transporters, targets or elements involved in the regulation of the expression of resistance determinants (as antibiotic inactivating enzymes or efflux pumps) had been searched under the idea that they should be the main cause of resistance. Based on this knowledge, specific applications such as the web tool PointFinder [80] able to identify known point mutations after mapping reads against reference genomes, as well as dedicated databases focusing on SNPs known to be involved in the acquisition of resistance in specific organisms have been developed [100]. However, efficient blind analysis of novel elements involved in the acquisition of resistance has been feasible just when WGS has become available. Massively parallel sequencing combined with traditional transposon mutagenesis [101], as well as high-throughput screening of transposon insertion libraries, have been used to track genes whose inactivation modifies the susceptibility to antibiotics [102,103]. This method has been used to analyse the resistome of different pathogens, including *Pseudomonas aeruginosa* [104–119], *E. coli* [110], *Staphylococcus aureus* [111,112] or *Klebsiella pneumoniae* [113], among other bacteria [114,115].

It is worth mentioning that the aforementioned studies only allow the identification of inactivating mutations that alter the susceptibility to antibiotics. However, AR is frequently achieved by mutations that do not inactivate the protein but just modify its activity. The use of adaptive laboratory evolution (ALE) experiments (Fig. 2) followed by WGS has allowed to describe AR mutations even before they are found in clinical settings. A major concern of these approaches is the differences between *in vivo* and *in vitro* growing conditions; it might be possible that *in vitro*-selected antibiotic resistant mutants are not selected *in vivo*, because fitness costs associated to the acquisition of resistance might be habitat-dependent. However, on several occasions, the same genetic events involved in the acquisition of AR *in vitro* have been also encountered in clinical isolates [116–119], validating the efficacy of these approaches.

Besides genomic studies, transcriptomic, proteomic and, eventually, metabolomic analyses have been implemented with the aim of deciphering novel mechanisms of resistance [122–125], or to detect the substrates, besides antibiotics, of known resistance determinants as multidrug efflux pumps [126]. The main drawback

of these approaches is that, although the acquisition of AR usually implies changes in the level of expression of several genes/proteins, most of these changes are not associated with the mechanism of resistance, but with the fitness costs due associated with such resistance (see below). In this respect, unless the expression of an already known resistance determinant changes, expression-based approaches alone do not always provide accurate information on the mechanisms of resistance; experimental validation is needed to distinguish between correlation and causality [127,128]. Further, it has been shown that antibiotics may induce an uncoordinated response in which transcriptionally and phenotypically important genes are not linked [129]. Nevertheless, this decoupling does not mean that transcriptomic/proteomic data are irrelevant for analysing AR, and there are different examples of their application [130–134]. Indeed, these methods have been used to describe new mechanisms of resistance, which have been experimentally validated [135], and it has been proposed that transcriptome profiling can be a useful tool in the analysis of AR in organisms such as *P. aeruginosa* [136].

In addition to the study of antibiotic resistant mutants, holistic expression-based studies (proteomics and transcriptomics), based on the analysis of the differential expression of some genes in presence of antibiotics, have been performed with the aim of describing novel mechanisms of resistance [137]. It is important to note that while it is true that transient resistance can be achieved via changes in gene expression, not all genes whose expression changes upon antibiotic stress are involved in the response to this injury [138–141]. Careful analyses, taking into consideration available information and experimental validation of predictions based on changes in gene expression, are needed in order to establish the relevance of the findings derived from these studies. Besides its potential application for discovering novel mechanisms of resistance, transcriptomic-based assays in the presence of antibiotics have been proposed as a useful approach for the fast diagnosis of AR. In a first step, conserved differences of gene expression between susceptible and resistant strains are searched. Afterward, a small set of genes (around 10) are chosen and their expression in presence of antibiotics measured using fast hybridization techniques that allow to determine resistance in just a few hours [142].

4. Whole-genome based diagnosis and epidemiology of antibiotic resistance

The use of -omic techniques has been certainly a milestone in any area of biology, AR included. Besides their contribution in basic laboratory-based analysis, they are revolutionizing clinical microbiology studies [143] in two different areas: molecular epidemiology and diagnostic procedures.

Concerning the first, the full understanding of all the elements driving the emergence of resistance requires global approaches so that WGS is increasingly replacing current PCR-based techniques (see above). Nowadays, WGS studies that include hundreds or even thousands of strains are not uncommon. Whole-genome MLST (wgMLST) or core genome MLST (cgMLST) approaches have allowed, without the need of a reference genome, the taxonomical classification of isolates based on more than 1000 genes instead of the few ones analysed in classical MLST studies [63–70]. Although the fundamentals of MLST and wgMLST are the same, wgMLST (and cgMLST) have a much higher discrimination power, allowing a much detailed, fine grained, description of the phylogenetic structures of bacterial populations 'from domain to strain' [69]. The linkage of this information with specific databases containing controlled reference genomes for diagnostic use and regulation as FDA-ARGOS [77], can help for studying in more detail the population structures of microorganisms with relevance for human health
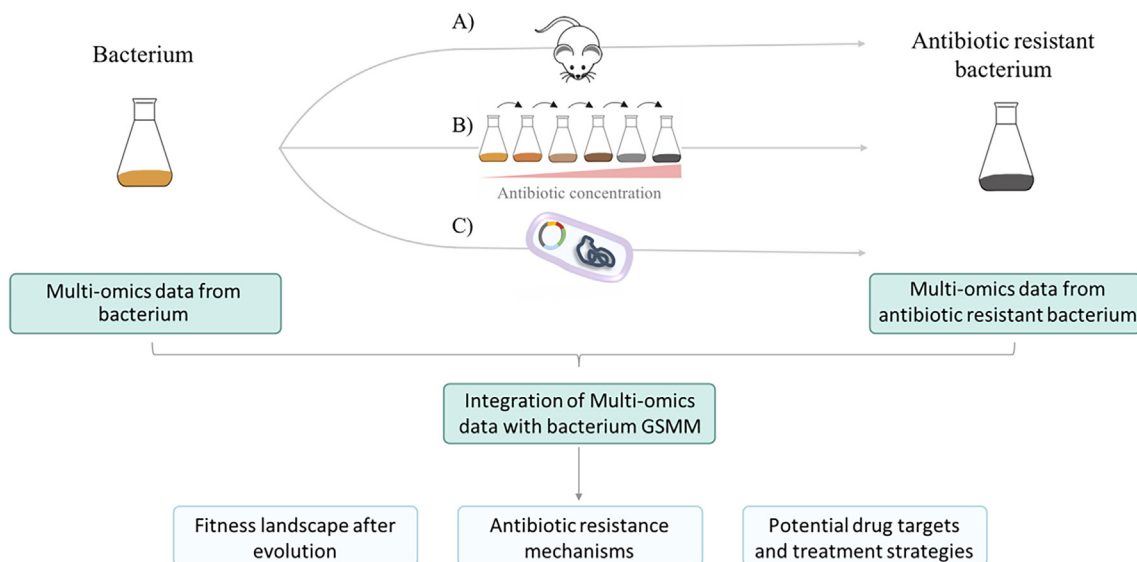
Fig. 2. Integrated research of AR by multi-omics data and pathogen-specific genome-scale metabolic models. Multi-omics data are collected from a parental wild-type strain and a derived and antibiotic-resistant bacterium. After that, pairwise comparison between both strains is performed. Resistance can be selected during infection upon *in vivo* evolution (a), from adaptive laboratory evolution linages, which can be explored at different time points along evolution (b), or through the introduction of an ARG in the wild-type strain (c). The comparison of isogenic bacteria, with different levels of resistance, and isolated from the same patient provides the most accurate information on the evolution of AR during treatment [120,121]. High throughput -omics data obtained from the analysis can be computationally integrated onto genome-scale metabolic models. Analyses of metabolic changes and mechanisms of AR allow the discovery of potential drug targets and treatment strategies against antibiotic-resistant pathogens. Fitness landscapes show the fitness cost of the AR, that influences the rate of development, stability and transmission of resistance, and the rate of resistance loss in the absence of the selecting antibiotic.

Concerning AR determinants, several approaches for finding ARGs and resistance mutations have been implemented (see above) and are in use for epidemiological studies. However, besides finding the resistance genes, it is also relevant to describe the mobile elements carrying them. It is important to note that, among all genomes deposited in databases, only a minor part of them is fully assembled; most are found as a variable number of different contigs, making in occasions difficult to distinguish if an ARG is present in a plasmid or in a chromosome. However, methods that allow getting long reads are helping to solve this problem; together with assemblers as SPAdes that include pipelines (plasmidSPAdes) for assembling plasmids [144] and informatic tools as PLACNET that allow the reconstruction of plasmids sequences obtained *de novo* or already present at WGS databases [145]. These tools can be used in order to analyse the genome/mobilome integration as well as to reconstruct the history of plasmid evolution and spread among different bacterial lineages [145]. More recently, a Perl application (Accessory genome Constellation Network-AcCNET) has been developed to compare the accessory genome (eventually containing ARGs) of several organisms, using the proteomes deduced from the analysed genomes [146]. Novel developments along this line as Pangenome Analysis Toolkit (PATO) allow to create, using conventional computers, clusters of thousands of genomes (pan-genome) using a list of pre-defined clusters, eventually from external sources as MLST; in addition, pangenomes can be filtered by genome frequency (i.e. the number of genomes that belong to the pangenome), and the frequency of ARGs [147].

While the methodologies for genome-based molecular epidemiology studies are already robust, a problem exists in databases, particularly in their reutilization for *meta*-analysis purposes. This problem consists of the general weakness (eventually absence) of metadata associated with the deposited sequences. Information regarding inclusion criteria, point of isolation, treatment, disease or age in the case of patients; chemical composition or potential pollution of the habitat in the case of isolates from natural ecosystems; and any other information that might influence emergence,

evolution and spread of AR is nearly absent and should be included in databases. In addition, databases are biased by the overrepresentation of strains derived from epidemic events. Although there are some databases that incorporate these metadata [148,149], their development is still in its infancy.

WGS has also been proposed as a good tool for simultaneously determine the phenotype of resistance and the mechanisms involved in this resistance. For this purpose, WGS databases that include in their metadata phenotypic information are valuable tools [150,151]. The use of machine-learning approaches able to predict AR based on sequence and expression data has been as well explored [152]. However, phenotypic tests are still superior to genome-based approaches, despite the utility of the latter in the case of slow-growing, non-cultivable or difficult-to-sample bacteria. One of these microorganisms is *Mycobacterium tuberculosis*, WGS has been proposed as a diagnostic tool that may be useful to analyse AR [153,154] and the *in vivo* evolution of resistance even prior to tuberculosis diagnosis [155], and dedicated databases comprising hundreds of AR mutations have been implemented. Although they are becoming affordable, time and resources required for genome analysis are still a problem for their use at clinical setting. An important drawback to the general application of these approaches as diagnostic tools is that they just provide predictions. Although genotypic-phenotypic linkage seems to be strong in the case of mutation driven resistance [156] and standalone informatic tools have been developed to link genotypic-phenotypic information [157], in the case of resistance genes the phenotype of resistance depends not only on the ARG, but also on its genomic context [24]. Further, sequence-based approaches cannot identify as yet unknown mechanisms of resistance that may contribute to the phenotype. Another drawback of WGS-based approaches is that they usually require the isolation of the organism. To avoid this problem, metagenomic analysis using techniques as Nanopore, which is portable, fast and renders long sequence reads, can be implemented at least in low-complexity microbiomes, as those involved in infections at orthopaedic implants [158]. Whole genome enriching methods targeting

specific organisms might be also applied in the case that the most probable causal agent of the disease is known. Indeed, these methods have been used for improving the recovery of *M. tuberculosis* DNA from sputum [159] in tuberculosis patients.

5. Antibiotic resistance beyond boundaries

The problem of AR goes beyond a single resistant bacterium or an infected patient. It is a global problem that involves the transfer of ARGs among bacteria forming part of gene-exchange communities [160], which are present in different ecosystems, including not only human-associated microbiomes, but also animal and environmental microbiomes [14,35,36,118,161–166]. Culture-based approaches just allow studying a minority of the populations present in a microbiome. Consequently, the implementation of metagenomic techniques, which do not require culture, as well as the development of tools for the analysis of metagenomes such as MEGAN [167] have been fundamental for a deep understanding of the taxonomic composition and the overall genetic capacities (including AR) of any microbiome.

AR is a global problem, in which any ecosystem (from natural environments to human hosts) may participate [168–170]. The socioeconomic issues that modulate this transmission [171], and the mathematical models addressing this aspect of the problem [172–175] will not be reviewed here. However, tools for tracking the transmission of ARGs among different ecosystems and geographical allocations [176], as well as detailed analyses of the effect of the selection of AR in a given microbiome are needed. In this regard, comparison of the resistomes of interconnected microbiomes [177] can provide clues of the transmission routes of resistance genes [35], and of the presence of these genes in bacteria colonizing ecological distinct niches as soil or human host [178]. Extremely randomized tree algorithm, optimized with Bayesian approaches, have been used to determine the variability of ARGs in different ecosystems and to track the ones that are discriminatory [179]. Besides, the analysis of the transmission of ARGs among ecosystems under strong selective pressure will certainly provide information on AR transmission. The microbiomes that are under stronger antibiotic selection are those from hospitalized humans and farms' animals. Accordingly, it has been stated that the use of antibiotics in farming may have important consequences for the spread of resistance in human populations [180–182]. ARGs selected in farms might be transferred to human pathogens or, eventually, the same pathogen can infect humans and other animals [183]. While this is generally true, there are some aspects of the problem that still require fine-tuning studies. Although, the species that infect humans and other animals, particularly in the case of mammals, can be the same, for some of them, the clonal lineages infecting each of the hosts are different [184]. Linking ARGs with the host is hence of relevance for estimating the risk to human health of such genes found in non-human-linked microbiomes. Besides, specific studies focusing on the mobilome shared by humans and other animals, may also help to analyse animal/human ARGs transmission [46,185,186].

Despite tools such as SqueezeMeta [187], a pipeline that integrates a set of programs required for metagenomic analysis, as well as the implementation of different AR databases (Table 1), are certainly helping in metagenomic studies on AR, some basic issues must be afforded when these studies are performed. The bioinformatic categorization of ARGs is one of them. This definition is usually carried out using DNA-based Blast searches oriented to already known resistance genes present in databases. In this regard, it is worth mentioning that existing ARG databases are skewed and biased towards beta-lactams resistance and aerobic/facultative aerobic bacteria, mainly those causing human infections. This bias

has to be taken into consideration when studies of the global resistome of a given habitat -particularly environmental ones- is performed [36]. In addition to *in-house* Blast search approaches, tools linked to specific databases as ARIBA [86], ARG-ANNOT [81], AMRFinder [188], MEGARes [189], CARD [82] or Resfinder [190] have been fruitfully used for finding putative resistance genes, and AR mutations, in a variety of genomes and metagenomes. Besides classical pairwise comparison approaches, probabilistic annotation algorithms, such as hidden Markov models, have demonstrated to be useful for detecting ARGs [79,191]. Although the predictive potential of these tools has largely improved, there are some issues that need to be taken into consideration when interpreting the results of the analyses of resistomes; in terms of the risk that detecting these putative resistance genes may have for human health [5,48]. The strength of the constraints used for *in silico* searches is one of them. In order to get a comprehensive view of the resistome, some authors use relaxed parameters for ARGs search. While this can be valid for creating a picture of the whole potential resistome, the results can be confusing to specifically track the resistance genes acquired by human pathogens after the discovery of antibiotics, which are the ones currently being a risk for human health [5].

The use of the above-mentioned approaches leads to the description of several ARGs in any studied metagenome [192–199]. However, most of them are unlikely to be a problem for human health because they are not present in human pathogens. Further, as discussed in [200], ARGs abundance does not necessarily correlates with risks to human health. Even more; the genomic context, which may drive the expression levels of a given ARG is fundamental for its contribution to the AR phenotype in its host [24]. It is true that all genes acquired by human pathogens were previously present in the genome of another microorganism [201], but the fact that the number of resistance genes present in the so far studied microbiomes is orders of magnitude above those currently moving among pathogenic bacteria, indicates that the probability of each of the detected genes to jump inside a human pathogen is low, unless it is already present in a mobile element [5]. In this regard, it has been shown that the resistome is linked to both ecology and phylogeny [202,203]. This is an indication that most detected ARGs belong to the intrinsic resistome, formed by the ensemble of genes present in all (or most) members of a given bacterial species, which contribute to its characteristic phenotype of antibiotic susceptibility and had not been recently acquired by HGT [103,104,204–207]. As discussed in [5], these genes have to be considered as phylogenetic markers, more than a risk of for human health. Specific databases on mobile elements as ACLAME [208], INTEGRALL [207], or mMGE (specifically focusing on human metagenomic mobile genetic elements) [209] are certainly helping in distinguishing between intrinsic and mobile ARGs.

While DNA-based approaches are useful for detecting already known ARGs, functional metagenomics [178,199,203,210,211] or structural based approaches can be used for predicting fully novel resistance genes [212]. Using the later, it has been established that most ARGs present in the human gut microbiome belong to the intrinsic resistome. Hence, more than being a risk for the acquisition of resistance by bacterial pathogens, they are resilience genes, able to reduce the perturbations due to antibiotic use and help to keep the homeostasis of human gut microbiome [213].

For the sake of comprehensibility, most AR databases contain intrinsic and acquired ARGs and, on several occasions, AR mutations are also included. When automatic analysis of these databases is performed, manual curation of the results is needed [204]. For instance, the finding of intrinsic resistance genes as *acrAB*, which encode an enterobacterial efflux pump [214], in a metagenome just means that *Enterobacteriaceae* are present in the microbiome. Similarly, the finding of genes as *marA* (encoding a global regulator) [215] or *gyrA*

(coding for the target of quinolones) [216,217] is not a sign of AR unless the mutant allele (that confers resistance) is found. Information on whether the cause of resistance is a mutation in the gene included in the database and not just to the presence of the gene, if the gene can confer by itself resistance or is a regulator of the expression of the actual ARG, as well as if the gene has been found in mobile elements or is just an intrinsic ARG, must be included in annotations to allow an accurate analysis of resistomes. Hierarchically organized databases, based on well-curated data, including information on resistance antibiotic class, resistance mechanism and gene mobility [218] can help solve this problem.

In the case of metagenomic samples, a particular problem concerns the identification of allelic forms of a resistance gene that differ just in a few nucleotides, as TEM-1-derived extended spectrum beta-lactamases [96]. Most metagenomic works are based in Illumina, short-reads, technologies. The assemble of these short reads may produce chimeric genes, not really present in an organism within the microbiome. The application of sequencing technologies allowing long reads seems to be a better approach to solve this problem [219]. While PacBio [220] is accurate enough to distinguish among alleles, Nanopore [221], and associated analytic tools as NanoGalaxy, which can be extremely useful to analyse arranges of ARGs, mobile elements [222,223], or the presence of acquired ARGs in a given microorganism (see below), are not accurate enough for the analysis of alleles presenting few differences in their sequences. Nevertheless, the portability and the fast generation of results of Nanopore devices make this technology an excellent choice for *in situ* metagenomic studies [224].

The risk for human health of an ARG depends on the organism in which this gene is present, ranging from environmental (lower) to commensals or pathogenic (higher). Binning approaches, based either on abundance or on nucleotide composition, are then needed for linking gene/bacterial host. [225,226], even rescuing potential novel species from metagenomic data (metagenomic species, see [227]). Applications such as GroopM [228], CONCOCT [229], SCIMM [230], Metawatt [231], MetaBAT [232], MaxBin [233], and others have been used for this purpose. However, although whole genome reconstruction from metagenomic data is feasible in low-complexity microbiomes, when a large number of sequence reads is available [234] or when the purpose is to analyse the pangenome of a given microorganism [235], these approaches present some limitations, particularly when commensal and pathogenic sublineages of the same species (as it happens in the case of *E. coli*) may co-exist in the same microbiome. The situation is even more difficult for analysing genomic islands and plasmids, presenting different copy numbers, with a nucleotide composition different to the one of the chromosome of the bacterial host, and eventually harboured by different hosts. In these cases, it has been reported that the percentage of mobile elements that are reconstructed and rightly binned to the chromosome of their bacterial host from metagenomic data is low [236]. Despite these problems, it has been recently stated that advanced sequencing techniques able to detect bacterium-specific methylation patterns allow a more accurate binning of chromosomes and plasmid sequences Further, non-bioinformatic, laboratory-based techniques, as EPIC-PCR [237]or Hi-C [238], may help to solve the problem of plasmid-chromosome binning.

A final concern refers to the abundance of ARGs reads in metagenomic data, since these reads can fall below the detection limits using regular metagenomic sequence depth. This is particularly relevant in the case of complex and/or environmental microbiomes, in which the resistant bacteria carrying the mobile ARGs constituting a risk for human health represent a very low fraction. Gene-capture systems, which allow targeted metagenomic studies, have been developed to enrich ARGs sequences in metagenomic analysis of complex microbiomes [239].

6. Natural computing applied to study hierarchically organized systems

As stated before, AR develops in the whole biosphere as a hierarchically organized complex system with several layers of organization, each one affecting the others. Its performance is based on homeostatic rules able to digest and integrate entropy (randomness) to keep identity, while allowing variation and evolution [240]. In Platonic terms, a cell is model of Nature; a biological microcosm reflecting the natural macrocosm. This has inspired the development of natural computing, nature-inspiring computational methods [241]. Among them, cellular computing includes approaches as artificial neuronal networks, or cellular automata, where ensembles of cells produce emergent qualities using communication strategies ("bacterial computing"), and bioinspired processors [242]. The latter includes network evolutionary processors transducers to simulate chemical transformations of substances or network genetic processors.

If the progress of genomic techniques has provided a powerful analytic tool to understand the biological complexity of the microbiosphere, new methods of synthesis are required to understand evolutionary functioning, including AR. Synthesis has benefited from the concept "cells as teachers", giving rise to new computational engineering methods that can be applied at higher levels [205,206].

One of the branches of cellular computing is membrane computing, taking advantage of the knowledge of the cellular organization to approach modelling of hierarchically structured biological systems. A cell is a space delimited by a membrane, which allows interaction with the environment while providing cellular individualization and identity; in fact, death essentially consists in the loss of individuality by membrane disruption [243]. Inside this space, there is a set of independent "objects", surrounded by membranes in the case of eukaryotic cells -as nucleus, mitochondria, or chloroplasts- or discrete self-assembled (in a sense, compartmentalized by a virtual membrane) entities -as ribosomes, or efflux pumps-; and in bacteria, independent self-replicative genetic elements as plasmids. This nested (tree-like) ontology of individuals inside individuals, membranes inside membranes, can be expanded downwards to other independent discrete units, as genes themselves, or associated ensembles of genes (operons). However, it can be also expanded upwards, as cells are individuals forming part ("inside") of individual supra-structures -as tissues, organs, species, microbiotas or holobionts- and finally ecosystems, which can be also treated as surrounded by functional membranes. Many of these intracellular or supracellular objects are independent units of selection [42,244,245].

The problem to address is how the evolution of a lower rank object influences the evolution of a higher rank object in the hierarchy, and vice versa. Indeed, the analytic (reductionistic) approach to study the complexity of natural systems and to ascertain the risks of their variation for human health has evident experimental boundaries. Empirical science is limited by the number of testable variables and by the practical impossibility of combining different ranges of them at various levels of the natural world. Mathematical modelling of AR has been limited until recently to "between-hosts" or "within-treated host" studies, but unfrequently using both dimensions [175]. Nevertheless, cellular computing allows the combination of both dimensions, at the same time that adds the "between-cells" and "within cells" dynamics of ARGs [246].

The series of "objects" defined by physical or virtual membranes in membrane computing models have interactive rules, integrating their action to sustain life. This view constitutes the basis of P-systems, one of the computational models used in mem-

brane computing, developed by Gheorghe Păun [247]. Membranes (including whole cells, or higher structures as microbiota) have rules by which they can replicate, propagate, dissolve, exchange information according to flexible rules, mutate, be selected by external agents, or be transferred into other membranes. For instance, an ARGs-containing plasmid moves from one cell to another, or a resistant clone moves from one microbiota to another, based on probabilistic events. Alterations in an object, as those derived from different mutation rates, are quantitatively reflected by differences in membrane fitness. Another kind of "objects" can be introduced (or excluded) from the system to influence the rules of the membrane-contained objects; for instance, certain concentrations of antibiotic might reduce the rate of replication or the rate of plasmid transfer between cells.

7. Membrane computing and the dynamics of plasmid-mediated lateral transfer of antibiotic resistance genes

The full analysis of the emergence, spread and evolution of AR requires to consider a highly complex multi-parametric and multi-hierarchical landscape [248]. Membrane computing allows not only to study the effect of changing each one of these parameters in the model and observe the effects this change causes, but also to combine, in a single computational run, the associate effects of changes in multiple parameters [249,250]. Tools as ARES [250] have been developed to analyse the *trans*-hierarchical evolution dynamics of AR in nested ecological compartments. Despite the intention of such tools is being user-friendly, their utilization still requires some bioinformatic knowledge, particularly on basic principles of membrane computing. The main advantages of these studies are: 1) the possibility of performing modelling *à la carte*, using the real data from a hospital, communities, or regions; 2) to use "range of possible parametric values" when some data are unavailable; 3) to evaluate the influence of a wide range of probabilities in the events influencing AR; 4) to assess the possible efficiency of interventions directed to reduce the burden of AR. In fact, membrane computing offers not only a way of representing processes but also a way to virtually reproduce (as a microworld) the process in a system.

Several factors influence the dynamics of AR plasmids transfer among bacterial populations. These factors have been analysed under *in vitro* conditions, but determining their relative weight in the dissemination and maintenance of AR in the real world is an impossible task using classical tools. Membrane computing is an appropriate method to cast light in this question, because of the nested nature of the biological units involved: genes into plasmids, plasmids into cells, bacterial cells among bacterial populations and microbiotas in a hospital or community (Fig. 3). The problem of plasmid-driven dissemination of AR has been addressed investigating the independent effect of these factors in the evolution of resistance, using the landscape described in the former paragraph [251].

This study has shown that plasmid conjugation frequencies significantly influence the spread of AR strains whose resistance determinants are encoded by the element only at transfer rates of 10^{-3}. Coexistence or exclusion (incompatibility) between plasmids with similar replication mechanisms was tested, and it has been described that coexistence might occur if the host strain contains two copies of these plasmids. Spontaneous plasmid loss has been found to have a small influence, being only of certain relevance at frequencies of loss of 10^{-4}-10^{-5}. Harbouring a plasmid might produce a reduction in fitness for the bacterial host; however, the model indicates that only reductions in growth rate above 6% influences the prevalence of plasmid carrying strains in a complex reality-like scenario. Plasmid fitness cost can be alleviated by "compensatory mutations", but the effect of such compensation is extremely low under the analysed conditions, except in strong hypermutable strains (10^{-5} of mutation frequency). All in all, this study [251] indicates that changes in evolutionary units located at sub-cellular scale might have some consequences in the "big-real" world constituted by higher hierarchical level units.

8. Membrane computing and the epidemiology of antibiotic resistance

Membrane computing models have been applied to understand the multilevel dynamics of antimicrobial resistance within a hospital [252]. The evolution of several different bacterial species and clones introduced at the beginning of the process at different abundances and presenting different AR phenotypes was studied. Given the time-lapse (see below) and the involvement of patients, this type of study cannot be experimentally addressed.

To get this goal, different species and strains were included in the model. *E. coli* was included with four initial phenotypes: fully susceptible, plasmid-mediated aminopenicillin-resistance, ciprofloxacin-resistance, and both aminopenicillin and ciprofloxacin-resistance; *K. pneumoniae*, which is intrinsically resistant to aminopenicillins, was plasmid-mediated cefotaxime-resistance and ciprofloxacin-resistance; *Enterococcus faecium* (intrinsically resistant to cefotaxime) was included with two phenotypes: aminopenicillin-susceptible and aminopenicillin-resistance (encoded in a conjugative, transmissible element); and *P. aeruginosa* (intrinsically resistant). The model allows to treat modelled individual patients with different antibiotics at different dosages and was implemented for mimicking 20,000 steps (one hour each), representing a total of around 2.3 years. New resistance phenotypes emerged during the modelization process, due to the transmission of ampicillin and cefotaxime from resistant to susceptible populations, and/or mutational events leading to ciprofloxacin-resistance. Notably, these resistance phenotypes are more frequent in hospital patients than in community ones, and the evolution of resistance depends on the basal R situation.

The study shows the importance of hospital admission-discharge rates in hospital; if this parameter decreases, resistance increases in the hospital. The reduction in hospital bacterial transmission (measured as the proportion of individuals that acquired any kind of bacteria from another individual per hour), also reduces the evolution of AR even though less than expected, probably because of the transmission of susceptible bacteria. The transmitted bacterial load was shown to facilitate colonization. More important than hospital transmission was the proportion of hospital patients treated with antibiotics, enhancing the proportion and complexity of AR. The effect of rapid or slow bactericidal effect of antibiotics was tested and it was shown that slow killing facilitates resistance. The importance of the intensity of the "ecological effect" of different antibiotics in decreasing susceptible populations of the gut, and therefore the colonization-resistance of the microbiota (favouring colonization, abundance, and transmission of resistant organisms), was also supported by the study. Finally, the strength of antibiotic selection on resistance traits was investigated; for instance, the possibility that ampicillin selects for clones harbouring ESBLs, even though this antibiotic is not the most clinically-relevant substrate of these antibiotic inactivating enzymes, was approached. This summary of results is just an example of the possibilities of membrane computing in predicting the evolution of AR at different levels and under disparate situations.
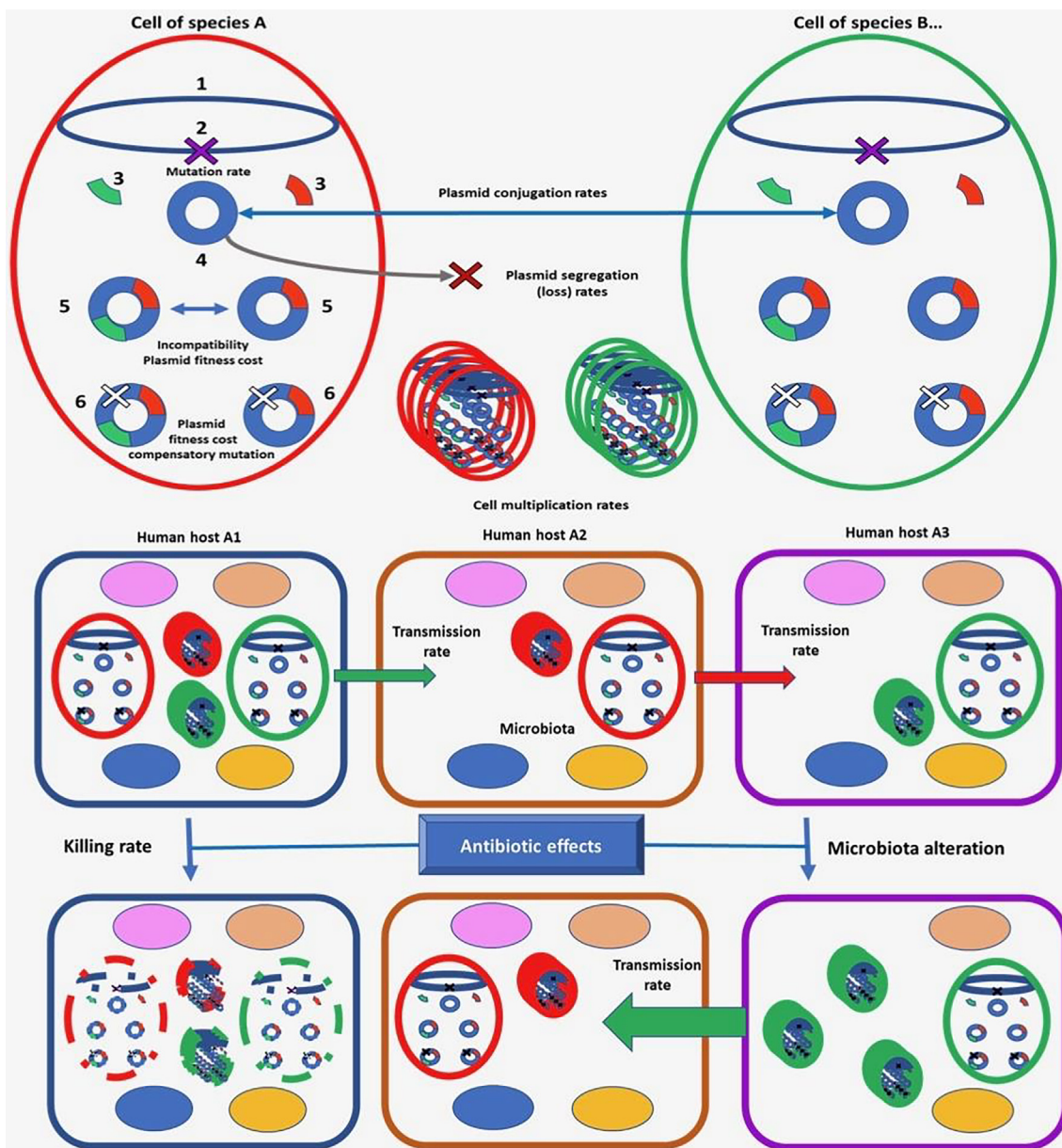
Fig. 3. Thought scheme of a membrane computing model to simulate the evolution of AR. The main factors influencing evolution of resistance are shown in the figure. Numbered parts inside the cell: 1) bacterial chromosome; 2) mutational event leading to AR; 3) ARGs; 4) mobile genetic element (plasmid); 5) compensatory mutation decreasing plasmid fitness cost for the cell. In the lower part, rounded squares represented different human hosts containing species (ovals) in the microbiota; bacteria are transmitted between hosts (green and red arrows, transmission of red or green strains). A given antibiotic can eliminate (lower, left) some types of bacterial cells (green, red) and their contents (broken lines) in host 1, and another one might also eliminate members of the microbiota in host 3 (lower, right), favouring the multiplication of resistant cells to this antibiotic (green), increasing the possibilities of being transmitted to untreated host 2. AR depends on "within and between cells" and "within and between hosts" evolution, in a complex nested system. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

9. Data integration for tackling antibiotic resistance

It has been generally assumed that the acquisition of AR produces a fitness cost that is mainly associated with an impaired bacterial growth. However, the situation is by far more complex; it is known that AR may alter the expression of several genes, not always associated with the acquisition of the resistance phenotype [23,253–256]. Besides altering the expression of virulence factors or signalling networks, the acquisition of AR may alter bacterial metabolism. These changes could be the Achilles' heel to specifically fight AR. Examples of metabolic rewiring associated with the acquisition of resistance are, among others, the anaerobic respiratory chain increased expression of *P. aeruginosa* multidrug resistant mutants [22,23], or the increasing expression of an rRNA

methylase in capreomycin resistant mycobacteria [257]. Knowing the effects of the acquisition of AR can hence provide evolution-based information for tackling this problem [258]. Conversely, the understanding of the bacterial metabolism can provide clues about the mechanisms and the effects of AR acquisition.

All the above-mentioned multi-omics technologies are increasingly used for the global analysis of AR. They simultaneously provide measurements of the expression of genes, proteins and metabolites that can be used to compare the physiology of susceptible and resistant pathogens at different time points [26]. However, information from these types of analysis is frequently disconnected, and the integration of this information may benefit from the establishment of cell models. The availability of an increasing amount of sequenced genomes and transcriptomes,

together with experimental work on the associated phenotypes, metabolic fluxes and biochemical studies, has allowed the reconstruction of genome-scale metabolic models (GSMMs) to elucidate the organisms' metabolisms and thus deduce metabolic capabilities for different species, several of them causing infections [26] (Fig. 2).

GSMMs interconnect metabolic reactions, their metabolites, genes and enzymes into mathematical models to study cellular mechanisms. Thereby, they predict metabolic responses linking gaps between transcriptional and phenotypic responses, a feature particularly relevant to understand the bacterial response to antibiotics [259–266]. In this regard, GSMMs together with multi-omics data are the first step to a quantitative analysis of pathogens physiology, including their metabolism, resistance and virulence mechanisms, hence constituting novel approaches for metabolism-based drug target studies [26].

An example of how -omic data can be used to determine metabolic changes in response to antibiotics is provided by the analysis of *P. aeruginosa* biochemical pathways perturbed by polymyxin treatment [260]. To get this information, the *P. aeruginosa* iPAO1 metabolic model together with RNA-Seq data was used. Apart from the known lipid A modifications associated with the response to polymyxin injury, numerous biochemical pathways involved in central, amino acid and fatty acid metabolism were perturbed, which enhances oxygen uptake and decreases growth rate [260,267]. Another example is the *in-silico* analysis of *S. aureus* small colony variants (SCVs). A GSMM of this organism showed that an *in silico*-generated *hemB* SCV mutant should have an essentially fermentative behaviour, should secrete lactate and its growth would be, at least partially, recovered in presence of glutamate, glutamine or arginine [28]. It remains to be established if this growth recovery means a reduction in AR of SCVs, in which case, these amino acids might be used as antibiotic adjuvants. Within the same study, the effect of *in silico* deletions of different genes on bacterial metabolism was analysed. Notably, all lethal gene deletions that had been identified experimentally were correctly predicted by the model, supporting its reliability, and allowing to propose enzymes from glycan biosynthesis and lipid metabolism as promising targets for the search of *S. aureus* inhibitors.

The construction of GSMMs is a tedious labour, in which four main stages – draft reconstruction, manual curation, conversion to mathematical model and network analysis – are needed to create a metabolic model [268,269]. There are different tools to create metabolic reconstructions [270–274] and among them, new automated tools, such as CarveMe universal model, allow the reconstruction of both species and community metabolisms, creating amazing metabolic model collections [275]. This automated tool overcomes the problems of previous methods, such as gap-filling, correction of elemental balance or directionality of reactions, detection of futile cycles and removal of blocked reactions or dead-end metabolites [275]. In addition, a flexible tool capable of combining information of GSMMs to identify promising targets as well as multi-target treatments is the web application FindTargetsWEB, which has been used to evaluate metabolic networks of both Gram-negative and Gram-positive multidrug-resistant pathogens such as *P. aeruginosa* or *S. aureus* [276].

Together with the use of -omic approaches, GSMMs across multiple strains and species or communities provide knowledge of pathogenic shifts due to AR at system-level [26] (Fig. 2). As an example, the use of GSMMs together with metabolomics data has been used to study streptomycin and chloramphenicol resistance in *Chromobacterium violaceum*. This study revealed that, while chloramphenicol resistance enhanced acetate production, streptomycin resistance increased acetate and formate secretion [259]. In addition to describe common metabolic pathways, GSMMs are essential to describe specific metabolic pathways for each microorganism, including AR ones, as well as virulence factor synthesis pathways [26]. This knowledge, together with transcriptomic, proteomic and metabolomic data can be used to understand the physiology of the bacterial pathogen under different conditions, including those with relevance for infection. For instance, this type of models has allowed to elucidate key elements associated to biofilm type of growth [277,278], to identify virulence factors [279] or to simulate the metabolic dynamic of *P. aeruginosa* in cystic fibrosis patients [280]. Further, recent works have identified common metabolic signatures associated with AR, virulence and clinical outcome in *P. aeruginosa* causing cystic fibrosis chronic infections [281] or metabolic profiles of multidrug resistant *M. tuberculosis* [282]. All this information is valuable for diagnosis/ prognosis purposes, as well as for finding pathways whose metabolic reprogramming may tackle resistance [283–287] or impair virulence aside from identifying metabolites that might be used for improving antibiotics' activity [284,288].

Even though transcriptomic/proteomic data have been used for target identification, the absence of correlation between transcriptional and phenotypical responses triggered by antibiotics calls into question if these studies could be generalized [129]. Transcriptional response related to antibiotics is unrelated to stress and confers no fitness advantage, thus a more useful approach to predict drug targets and eradicate pathogens would include network topological analyses to put into realistic models high-throughput experiments [129]. It is for these reasons that the development of GSMMs linking the metabolic phenotype to environmental and genetic perturbations is central to drug discovery [275,289].

There are different computational methods used to test and validate GSMMs in the field of drug target discovery. Firstly, flux balance analysis (FBA) is used to represent the possible behaviour of microbial metabolism and to predict gene essentiality [290]. One of the objectives of FBA is the use of biomass production studies to detect new elements able to compromise pathogen's growth. This approach has provided insights into *Porphyromonas gingivalis* essential growth and has postulated LPS production as a potential drug target [291]. Moreover, FBA is used to determine if the deletion of a given gene can arrest the selected function in the bacterial metabolism, being the gene knockout stimulation using FBA the most popular method employed for drug discovery [289,276]. Secondly, flux variability analysis (FVA) determines the range of alternate routes able to get the same objective, identifying other potential drug targets [292]. Thirdly, flux sampling calculates all solutions in a statistical meaningful way when the objective is not clear [293].

Traditionally, the most common method to identify drug targets has been the prediction of essential genes whose inactivation leads to flux redistribution, being the enzymes they encode potential drug targets. GSMMs are highly effective in gene deletion analysis studies, since this type of computational analyses, performed in a few seconds, save time and effort in comparison to experimental work [26]. Consequently, GSMMs have allowed to identify essential enzymes that can be exploited as targets of broad-spectrum antibiotics against different pathogens. For instance, these studies have revealed that *Yersinia pestis* essential genes are mainly genes encoding proteins involved in amino acid transport and metabolism and that some of these metabolic genes are also identified as essential in *E. coli* and *Salmonella typhimurium* [294]. The use of GSMMs for the identification of the organisms' catabolic cores might then be a useful tool for identifying pathways predicted to be critical for biomass generation, being even essential. The experimental validation of such prediction shows that this approach is indeed useful [295].

Besides the identification of essential pathways, another approach is the prediction of essential metabolites [296]; when a metabolite is removed within the model, all reactions in which it

is involved disappear and the consequences are studied. The enzymes found to be involved in the consumption and production of these metabolites can be studied as new drug targets [263,289], and the information on the metabolites themselves can be used to search for analogues of them that might inhibit bacterial growth [289]. These approaches can also be used to explore the chemical space that may be eventually useful to develop combination antibiotic therapies [297] or multi-target antibiotics. For the later, sets of reactions that act together may be targeted as an unique pathway [298] and consequently, altering the flux of one of them by targeting one enzyme results in an altered flux of all the reactions. For instance, 147 reactions were found to be coupled in *M. tuberculosis* models and 25 of these reactions account with previously known drug targets; thus, lethal combinations of these reactions could be assessed as potential targets [299].

GSMMs may also allow to simulate different infection scenarios, predict cellular phenotypes by changing nutrient sources and deduce infection features in different infected locations. This information on the metabolic situation of bacteria at different stages of infection might serve for identifying specific biomarkers and putative drug targets not so easily identified through *in vitro* approaches. For instance, pathogens may require specific nutrients and resources from the host cells. One example of this situation is *Leishmania major*, which obtains certain amino acids from the host macrophages since it is unable to synthesize them. The inhibition of this kind of pathways associated to infective situations and host characteristics, which may lead to environment perturbations, may help in defining new targets and therapeutic options based in metabolism remodelling [300,301]. Among others, fatty acid and lipid metabolic reactions have been recently proposed as promising antimicrobial targets [26].

All in all, conventional methods to detect possible targets are expensive and slow. Accordingly, computational studies are a more efficient alternative to detect essential genes and their products as possible drug targets, facilitating, upon validation of the generated predictions, the development of new drugs targeting at different stages of infection [26]. Despite the fact that metabolic modelling is a new approach, it has allowed to detect targets that are consistent with previous data [28] as well as to provide new hypotheses [296] to explore.

Once drug targets are identified, drugs capable of inhibiting these targets should be screened. The most evident category for developing new drugs interfering with bacterial metabolism would be the structural analogues of the essential metabolites found by metabolic reconstruction and reprogramming through GSMM. These analogues would be able to replace and to compete with the essential metabolite, disrupting metabolism and causing cell death [289]. For instance, lysine analogues that target lysine riboswitches inhibit *Bacillus subtilis* growth by repressing lysine biosynthesis pathway. Some of these riboswitches regulate essential genes whose proteins produce metabolites that cannot be obtained from the external medium, being promising targets for the development of new RNA-binding antibiotics [302]. In addition to the search of new antibacterial compounds, these approaches could be used to identify metabolite analogues able to re-sensitize resistant pathogens to existing antibiotics. As mentioned before, *C. violaceum* resistance is induced by metabolic shifts, more specifically by a switch to fermentative metabolism thanks to an increased NADH/NAD ratio. GSMM studies have allowed to identify 2-oxoadipate as a compound able to re-sensitize these resistant populations, inducing susceptibility to existing antibiotics and death in *C. violaceum* [259]. Nevertheless, and despite these advances in the field, no novel antibiotic based on these approximations is in the market yet.

10. Summary and outlook

Before the worldwide COVID pandemic emerged, AR was considered as the most relevant One Heath, Global Health problem that might compromise human health and welfare in the next years [303,304]. Traditionally, AR studies have focused on infected patients. However, it is now clear that the problem goes outside hospitals and that AR evolution and spread involves a variety of ecosystems that, besides humans, include other animals (farm animals, pets and also wild-animals) and natural, non-clinical, ecosystems [36,200]. Most common methods to analyse the genetic and molecular bases of AR are gene-centric: the presence of a specific gene or mutation is sought and, eventually, the association of this gene with mobile genetic elements is determined. While relevant for detecting specific mechanisms of resistance, these approaches are not sufficient to provide a global picture of the elements involved in the emergence and spread of AR. In this regard, it is relevant to state that AR is a multihierachical process with different interconnected elements (from genes to global population) involved and the simultaneous analysis of these processes requires the implementation of holistic approaches. WGS and other -omic techniques (transcriptomics, proteomics and metabolomics) have been implemented in the last years as the most appropriate tools for the integrative analysis of AR. These tools have already demonstrated to be valuable in basic studies on AR and in molecular epidemiology studies. However, they still present some limitations in their use as diagnostic tools (see below). Further, several of the tools require some bioinformatic training by users. This is not necessarily a drawback if the teams include bioinformaticians and experts in AR, but the fast translation of the results into clinics requires the development of user-friendly tools, not always available.

There two areas that still require novel tools and information; one is the accurate prediction of AR based in WGS data, which is needed for implementing sequence-based diagnostic approaches. Mutation-driven resistance is likely an affordable task, although it still requires a more comprehensive information of all mutations leading to AR resistance in the different organisms under study. Indeed, recent studies on experimental evolution in presence of antibiotics have shown that there are still hidden AR mutations that should be incorporated into the models [116,305–307]. Comprehensive, organism-specific databases as MUBII-TB-DB, developed for predicting WGS-based *M. tuberculosis* AR have shown to be effective for this type of analysis [100]. The situation concerning acquired ARGs is more complex. Although there exist specific databases on mobile elements as ACLAME [208], INTEGRALL [207], or mMGE [209], most databases contain information on mutations, intrinsic and acquired ARGs. This information is fundamental for manual annotation of the results, but is less amenable for the automatic analysis of ARGs. Hierarchically organized databases as HMD-ARG [218] may help in solving this problem. However, even in these cases further refinement in the annotation is needed to distinguish between phenotypically validated ARGs and predicted ones. Rankings as those stablished in some genomic databases, as the one of *P. aeruginosa*, distinguishing among genes, which function has been validated in the organism under study, genes that are orthologs to genes with a validated function in another organism and genes with a function inferred just by homology [308] can help to distinguish between functionally validated and predicted AR, a feature with relevance for risks analysis [5,48].

The other area requiring further development is the binning between the identified ARGs and the organism carrying them (commensal or pathogen) when metagenomes are studied. The use of powerful, automatic analytical tools as PhyloPhlAn 3.0 [309], which uses hundreds of thousands of genomes and tens of

thousands of metagenomes for genome reconstruction may help in this endeavour. However, even in this case, the incorporation of HGT-acquired genes into the genomes is still cumbersome, unless the genomes containing ARGs are highly represented in the metagenome, because they present a different nucleotide composition and, if present in plasmid a different abundance. Sequence methods that provide long reads can help to solve this problem, and the combination of long- and short-reads information for assembling seems to be superior than using independently either short-reads or long-reads based methodologies, however costs will be higher too [310]. Hybrid methods making use of information on abundance and composition are the best option for binning the generated contigs [311]. However, even in this case, plasmid-chromosome binning can be cumbersome. The above-mentioned analysis of DNA methylation patterns that are specific of restriction/modification systems may help in these studies [312].

Once omic-based global information on AR is becoming accessible, time has come to implement models for its synthesis. The future perspective in the advancement of knowledge on complex biological systems will be highly dependent on the progresses in the integration of throughput analytic methods (-omics) and powerful synthetic tools (as natural multihierarchical computing, GSMMs, Big Data, machine-learning and metadata management, or artificial intelligence for systems engineering). In the case of AR, both detailed and global genomic information should become increasingly available, feeding integrative models to extract the predominant trends and to associate them with anthropogenic interventions. The information will constitute an important step forward for stablishing rational, ecological, and evolutionary approaches to tackle this relevant health problem.

CRediT authorship contribution statement

Teresa Gil-Gil: Writing - review & editing. **Luz Edith Ochoa-Sánchez:** Writing - review & editing. **Fernando Baquero:** Writing - review & editing. **José Luis Martínez:** Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

References

[1] WHO. Antimicrobial resistance: global report on surveillance. Geneva: World Health Organization; 2014.
[2] WHO. !!
[245] Baquero F. Genetic hyper-codes and multidimensional Darwinism: replication modes and codes in evolutionary individuals of the bacterial world. In: Trueba G, editor. Why does Evolution Matter? The Importance of Understanding Evolution. Newcastle upon Tyne. United Kingdom: Cambridge Scholar Publishing; 2014. p. 165–80.
[246] Jiménez MJP, Jiménez ÁR, Caparrini FS. Complexity classes in models of cellular computing with membranes. Nat Comput 2003;2(3):265–85.
[247] Păun G. Computing with Membranes. J Comput Syst Sci 2000;61(1):108–43.
[248] Martinez JL, Baquero F. Emergence and spread of antibiotic resistance: setting a parameter space. Ups J Med Sci. 2014;119(2):68–77.
[249] Baquero F, Campos M, Llorens C, Sempere JM. A Model of Antibiotic Resistance Evolution Dynamics Through P Systems with Active Membranes and Communication Rules. In: Graciani C, Riscos-Núñez A, Păun G, Rozenberg G, Salomaa A, editors. Enjoying Natural Computing. Lecture Notes in Computer Science. Switzerland: Springer, Cham; 2018. p. 33–44.
[250] Campos M, Llorens C, Sempere JM, Futami R, Rodriguez I, Carrasco P, et al. A membrane computing simulator of trans-hierarchical antibiotic resistance evolution dynamics in nested ecological compartments (ARES). Biol Direct 2015;10:41.
[251] Campos M, San Millán Á, Sempere JM, Lanza VF, Coque TM, Llorens C, et al. Simulating the Influence of Conjugative-Plasmid Kinetic Values on the Multilevel Dynamics of Antimicrobial Resistance in a Membrane Computing Model. Antimicrob Agents Chemother 2020;64(8):e00593–e620.
[252] Campos M, Capilla R, Naya F, Futami R, Coque T, Moya A, et al. Simulating Multilevel Dynamics of Antimicrobial Resistance in a Membrane Computing Model. MBio 2019;10(1).
[253] Olivares J, Alvarez-Ortega C, Linares JF, Rojo F, Köhler T, Martínez JL. Overproduction of the multidrug efflux pump MexEF-OprN does not impair Pseudomonas aeruginosa fitness in competition tests, but produces specific changes in bacterial regulatory networks. Environ Microbiol 2012;14:1968–81.
[254] Olivares J, Álvarez-Ortega C, Martínez JL. Metabolic compensation of fitness costs associated with overexpression of the multidrug efflux pump MexEF-OprN in Pseudomonas aeruginosa. Antimicrob Agents Chemother 2014;58:3904–13.
[255] Alcalde-Rico M, Olivares-Pacheco J, Alvarez-Ortega C, Camara M, Martinez JL. Role of the Multidrug Resistance Efflux Pump MexCD-OprJ in the Pseudomonas aeruginosa Quorum Sensing Response. Front Microbiol 2018;9:2752.
[256] Alcalde-Rico M, Olivares-Pacheco J, Halliday N, Cámara M, Martínez JL. The analysis of the role of MexAB-OprM on quorum sensing homeostasis shows that the apparent redundancy of Pseudomonas aeruginosa multidrug efflux pumps allows keeping the robustness and the plasticity of this intercellular signaling network. bioRxiv. 2020:2020.03.10.986737.
[257] Freihofer P, Akbergenov R, Teo Y, Juskeviciene R, Andersson DI, Bottger EC. Nonmutational compensation of the fitness cost of antibiotic resistance in mycobacteria by overexpression of tlyA rRNA methylase. RNA 2016;22 (12):1836–43.
[258] Baquero F, Martinez JL. Interventions on Metabolism: Making Antibiotic-Susceptible Bacteria. mBio 2017;8(6).
[259] Banerjee D, Raghunathan A. Constraints-based analysis identifies NAD+ recycling through metabolic reprogramming in antibiotic resistant Chromobacterium violaceum. PLoS One 2019;14(1):e0210008.
[260] Zhu Y, Czauderna T, Zhao J, Klapperstueck M, Maifiah MHM, Han ML, et al. Genome-scale metabolic modeling of responses to polymyxins in Pseudomonas aeruginosa. GigaScience 2018;7(4).
[261] O'Brien EJ, Monk JM, Palsson BO. Using Genome-scale Models to Predict Biological Capabilities. Cell 2015;161(5):971–87.
[262] Bosi E, Monk JM, Aziz RK, Fondi M, Nizet V, Palsson BO. Comparative genome-scale modelling of Staphylococcus aureus strains identifies strain-specific metabolic capabilities linked to pathogenicity. PNAS 2016;113(26):E3801–9.
[263] Kim HU, Kim SY, Jeong H, Kim TY, Kim JJ, Choy HE, et al. Integrative genome-scale metabolic analysis of Vibrio vulnificus for drug targeting and discovery. Mol Syst Biol 2011;7:460.
[264] Krueger AS, Munck C, Dantas G, Church GM, Galagan J, Lehar J, et al. Simulating Serial-Target Antibacterial Drug Synergies Using Flux Balance Analysis. PLoS One 2016;11(1):e0147651.
[265] Aziz RK, Monk JM, Lewis RM, In Loh S, Mishra A, Abhay Nagle A, et al. Systems biology-guided identification of synthetic lethal gene pairs and its potential use to discover antibiotic combinations. Sci Rep 2015;5:16025.
[266] Bordbar A, Monk JM, King ZA, Palsson BO. Constraint-based models predict metabolic and associated cellular functions. Nat Rev Genet 2014;15 (2):107–20.
[267] Hussein M, Han ML, Zhu Y, Zhou Q, Lin YW, Hancock REW, et al. Metabolomics Study of the Synergistic Killing of Polymyxin B in Combination with Amikacin against Polymyxin-Susceptible and -Resistant Pseudomonas aeruginosa. Antimicrob Agents Chemother 2019;64(1).
[268] Thiele I, Palsson BO. A protocol for generating a high-quality genome-scale metabolic reconstruction. Nat Protoc 2010;5(1):93–121.
[269] Norsigian CJ, Fang X, Seif Y, Monk JM, Palsson BO. A workflow for generating multi-strain genome-scale metabolic models of prokaryotes. Nat Protoc 2020;15(1):1–14.
[270] Agren R, Liu L, Shoaie S, Vongsangnak W, Nookaew I, Nielsen J. The RAVEN toolbox and its use for generating a genome-scale metabolic model for Penicillium chrysogenum. PLoS Comput Biol 2013;9(3):e1002980.
[271] Dias O, Rocha M, Ferreira EC, Rocha I. Reconstructing High-Quality Large-Scale Metabolic Models with merlin. Methods Mol Biol 2018;1716:1–36.

[272] Henry CS, DeJongh M, Best AA, Frybarger PM, Linsay B, Stevens RL. High-throughput generation, optimization and analysis of genome-scale metabolic models. Nat Biotechnol 2010;28(9):977–82.

[273] Karp PD, Latendresse M, Paley SM, Krummenacker M, Ong QD, Billington R, et al. Pathway Tools version 19.0 update: software for pathway/genome informatics and systems biology. Brief Bioinform. 2016;17(5):877–90.

[274] Pitkanen E, Jouhten P, Hou J, Syed MF, Blomberg P, Kludas J, et al. Comparative genome-scale reconstruction of gapless metabolic networks for present and ancestral species. PLoS Comput Biol 2014;10(2):e1003465.

[275] Machado D, Andrejev S, Tramontano M, Patil KR. Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. Nucleic Acids Res 2018;46(15):7542–53.

[276] Merigueti TC, Carneiro MW, Carvalho-Assef APD, Silva-Jr FP, da Silva FAB. FindTargetsWEB: A User-Friendly Tool for Identification of Potential Therapeutic Targets in Metabolic Networks of Bacteria. Front Genet 2019;10:633.

[277] Biggs MB, Papin JA. Novel multiscale modeling tool applied to Pseudomonas aeruginosa biofilm formation. PLoS One 2013;8(10):e78011.

[278] Vital-Lopez FG, Reifman J, Wallqvist A. Biofilm Formation Mechanisms of Pseudomonas aeruginosa Predicted via Genome-Scale Kinetic Models of Bacterial Metabolism. PLoS Comput Biol 2015;11(10):e1004452.

[279] Bartell JA, Blazier AS, Yen P, Thogersen JC, Jelsbak L, Goldberg JB, et al. Reconstruction of the metabolic network of Pseudomonas aeruginosa to interrogate virulence factor synthesis. Nat Commun 2017;8:14631.

[280] Oberhardt MA, Goldberg JB, Hogardt M, Papin JA. Metabolic network analysis of *Pseudomonas aeruginosa* during chronic cystic fibrosis lung infection. J Bacteriol 2010;192(20):5534–48.

[281] Moyne O, Castelli F, Bicout DJ, Boccard J, Camara B, Cournoyer B, et al. Metabotypes of Pseudomonas aeruginosa Correlate with Antibiotic Resistance. Virulence and Clinical Outcome in Cystic Fibrosis Chronic Infections. Metabolites. 2021;11(2).

[282] Rêgo AM, Alves da Silva D, Ferreira NV, de Pina LC, Evaristo JAM, Caprini Evaristo GP, et al. Metabolic profiles of multidrug resistant and extensively drug resistant Mycobacterium tuberculosis unveiled by metabolomics. Tuberculosis (Edinburgh, Scotland) 2020;126:102043.

[283] Vestergaard M, Nohr-Meldgaard K, Bojer MS, Krogsgard Nielsen C, Meyer RL, Slavetinsky C, et al. Inhibition of the ATP Synthase Eliminates the Intrinsic Resistance of Staphylococcus aureus towards Polymyxins. mBio 2017;8(5).

[284] Jiang M, Kuang SF, Lai SS, Zhang S, Yang J, Peng B, et al. Na(+)-NQR Confers Aminoglycoside Resistance via the Regulation of l-Alanine Metabolism. mBio 2020;11(6).

[285] Su YB, Kuang SF, Peng XX, Li H. The depressed P cycle contributes to the acquisition of ampicillin resistance in Edwardsiella piscicida. J Proteomics 2020;212:103562.

[286] Gardner SG, Marshall DD, Daum RS, Powers R, Somerville GA. Metabolic Mitigation of Staphylococcus aureus Vancomycin Intermediate-Level Susceptibility. Antimicrob Agents Chemother 2018;62(1).

[287] Zampieri M, Zimmermann M, Claassen M, Sauer U. Nontargeted Metabolomics Reveals the Multilevel Response to Antibiotic Perturbations. Cell Rep 2017;19(6):1214–28.

[288] Ye JZ, Lin XM, Cheng ZX, Su YB, Li WX, Ali FM, et al. Identification and efficacy of glycine, serine and threonine metabolism in potentiating kanamycin-mediated killing of Edwardsiella piscicida. J Proteomics 2018;183:34–44.

[289] Kim TY, Kim HU, Lee SY. Metabolite-centric approaches for the discovery of antibacterials using genome-scale metabolic networks. Metab Eng 2010;12(2):105–11.

[290] Kauffman KJ, Prakash P, Edwards JS. Advances in flux balance analysis. Curr Opin Biotechnol 2003;14(5):491–6.

[291] Mazumdar V, Snitkin ES, Amar S, Segre D. Metabolic network model of a human oral pathogen. J Bacteriol 2009;191(1):74–90.

[292] Mahadevan R, Schilling CH. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. Metab Eng 2003;5(4):264–76.

[293] Price ND, Schellenberger J, Palsson BO. Uniform sampling of steady-state flux spaces: means to design experiments and to interpret enzymopathies. Biophys J 2004;87(4):2172–86.

[294] Charusanti P, Chauhan S, McAteer K, Lerman JA, Hyduke DR, Motin VL, et al. An experimentally-supported genome-scale metabolic network reconstruction for Yersinia pestis CO92. BMC Syst Biol 2011;5:163.

[295] Hartman HB, Fell DA, Rossell S, Jensen PR, Woodward MJ, Thorndahl L, et al. Identification of potential drug targets in Salmonella enterica sv. Typhimurium using metabolic modelling and experimental validation. Microbiology (Reading) 2014;160(Pt 6):1252–66.

[296] Chavali AK, D'Auria KM, Hewlett EL, Pearson RD, Papin JA. A metabolic network approach for the identification and prioritization of antimicrobial drug targets. Trends Microbiol 2012;20(3):113–23.

[297] Campos AI, Zampieri M. Metabolomics-Driven Exploration of the Chemical Drug Space to Predict Combination Antimicrobial Therapies. Molecular cell 2019;74(6). 1291-303.e6.

[298] Papin JA, Reed JL, Palsson BO. Hierarchical thinking in network biology: the unbiased modularization of biochemical networks. Trends Biochem Sci 2004;29(12):641–7.

[299] Jamshidi N, Palsson BO. Investigating the metabolic capabilities of Mycobacterium tuberculosis H37Rv using the in silico strain iNJ661 and proposing alternative drug targets. BMC Syst Biol 2007;1:26.

[300] McConville MJ, de Souza D, Saunders E, Likic VA, Naderer T. Living in a phagolysosome; metabolism of Leishmania amastigotes. Trends Parasitol. 2007;23(8):368–75.

[301] Schwegmann A, Brombacher F. Host-directed drug targeting of factors hijacked by pathogens. Sci Signal. 2008;1(29):re8..

[302] Blount KF, Wang JX, Lim J, Sudarsan N, Breaker RR. Antibacterial lysine analogs that target lysine riboswitches. Nat Chem Biol 2007;3(1):44–9.

[303] Laxminarayan R, Van Boeckel T, Frost I, Kariuki S, Khan EA, Limmathurotsakul D, et al. The Lancet Infectious Diseases Commission on antimicrobial resistance: 6 years later. Lancet Infect Dis 2020;20(4):e51–60.

[304] Robinson TP, Bu DP, Carrique-Mas J, Fevre EM, Gilbert M, Grace D, et al. Antibiotic resistance is the quintessential One Health issue. Trans R Soc Trop Med Hyg 2016;110(7):377–80.

[305] Sanz-Garcia F, Hernando-Amado S, Martinez JL. Mutation-driven evolution of Pseudomonas aeruginosa in the presence of either ceftazidime or ceftazidime/avibactam. Antimicrob Agents Chemother 2018.

[306] Blanco P, Corona F, Martinez JL. Involvement of the RND efflux pump transporter SmeH in the acquisition of resistance to ceftazidime in Stenotrophomonas maltophilia. Sci Rep 2019;9(1):4917.

[307] Blanco P, Corona F, Martinez JL. Mechanisms and phenotypic consequences of acquisition of tigecycline resistance by Stenotrophomonas maltophilia. J Antimicrob Chemother 2019.

[308] Winsor GL, Griffiths EJ, Lo R, Dhillon BK, Shay JA, Brinkman FS. Enhanced annotations and features for comparing thousands of Pseudomonas genomes in the Pseudomonas genome database. Nucleic Acids Res 2016;44(D1):D646–53.

[309] Asnicar F, Thomas AM, Beghini F, Mengoni C, Manara S, Manghi P, et al. Precise phylogenetic analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0. Nat Commun 2020;11(1):2500..

[310] Brown CL, Keenum IM, Dai D, Zhang L, Vikesland PJ, Pruden A. Critical evaluation of short, long, and hybrid assembly for contextual analysis of antibiotic resistance genes in complex environmental metagenomes. Sci Rep 2021;11(1):3753.

[311] Borderes M, Gasc C, Prestat E, Galvão Ferrarini M, Vinga S, Boucinha L, et al. A comprehensive evaluation of binning methods to recover human gut microbial species from a non-redundant reference gene catalog. NAR Genom Bioinform. 2021;3(1). lqab009.

[312] Beaulaurier J, Zhu S, Deikus G, Mogno I, Zhang XS, Davis-Richardson A, et al. Metagenomic binning and association of plasmids with bacterial host genomes using DNA methylation. Nat Biotechnol 2018;36(1):61–9.