



Evolution and Transfer of Antibiotic Resistance

**FRONT PAGE** 

# **PROJECT FINAL REPORT**

**Grant Agreement number: 282004 Project acronym: EVOTAR Project title: Evolution and Transfer of Antibiotic Resistance Funding Scheme: Collaborative project Date of latest version of Annex I against which the assessment will be made: 21-08-2012 Period covered: from 1 October 2011 to 30 September 2015 Dr. Rob J. L. Willems, Associate Professsor, Universitair Medische Centrum Utrecht Tel: +31887557630 Fax: +31302541770 E-mail: r.willems@umcutrecht.nl**

**Project website<sup>1</sup> address: www.evotar.eu**

 $1/2$  The home page of the website should contain the generic European flag and the FP7 logo which are available in electronic format at the Europa website (logo of the European flag: http://europa.eu/abc/symbols/emblem/index\_en.htm logo of the 7th FP: http://ec.europa.eu/research/fp7/index\_en.cfm?pg=logos). The area of activity of the project should also be mentioned.

# *4.1 Final publishable summary report*

# **Executive Summary**

In the EvoTAR project the human gut microbiome was primarily used as a starting point for the study of antibiotic resistance. The gut harbours the most complex and abundant microbiota of our body and here antibiotic-resistance may emergence readily through selection of pre-existing resistant bacteria and gene transfer events for which there are ample opportunities in this environment. To study the dynamics of the resistome (*i.e.* the total antimicrobial resistance determinants (ARD) in the human gut the consortium exploited different culture independent technologies, such as full metagenomic sequencing, functional metagenomic selections and resistance genes capture platforms as well as novel high throughput culture methods. Metagenomic sequencing revealed that a relatively short-term exposure altered both the richness (a decrease of 67.8%) and the ARD abundance (a decrease of 65.2%) greatly. There was thus no enrichment of the ARDs, presumably because of the inability of most species to withstand the harsh antibiotic treatment. In contrast, a long-term (chronic) exposure altered the relation between the richness of the microbiome, which was decreased and the abundance of the ARDs, which was increased. Clearly, such exposure selects for the species than can strive in constant antibiotic presence, due to the ARDs they encode. Also functional selections showed that hospitalization and antibiotic treatment has profound effects on the gut resistome, with a vast expansion of the resistome in some patients. This expansion of the resistome during hospitalization could lead to an increased risk of transfer of antibiotic resistance genes to infecting pathogens. Importantly, preliminary data indicate that 6 months after discharge, the abundance of antibiotic resistance genes return to the same level. Also with the newly developed gene capture platform (described below) important changes in the composition of antibiotic-resistance genes in samples form hospitalized patients were observed, with remarkable gains and loss in the recovery of certain families of genes.

In addition to the culture-independent methods optimized cultivation methods for the human gut microbiota were developed that can capture a representative majority of the cells present in a sample by both abundance and overall community structure. These novel culture technologies in combination with whole genome sequencing revealed important reservoirs of antibiotic resistant organisms and antibiotic resistance genes in soil and marine environments. One example was the identification of 17 new carbapenemases with 30-76% amino acid identity to previously confirmed carbapenemases. Our results show that the environment including soil and water is a source of highly diverse carbapenemases that are produced by a variety of bacterial species and have not yet emerged in clinical settings. These carbapenemases may constitute potential carbapenem-resistance determinants of clinical relevance if acquired by pathogenic bacteria, as they were functionally expressed in *E. coli*.

Dedicated research on the evolution and spread of resistance in Enterobacteriaceae and enterococci revealed, among others, low-likelihood of recent transmission of ESBL-*E. coli* between poultry and humans but that virtually identical ESBL-carrying plasmids were shared by genetically unrelated human and poultry isolates, strongly suggesting that ESBLs are mainly disseminated via epidemic plasmids that can spread between different reservoirs. The latter was demonstrated using the PLACNET tool mentioned below.

A major objective of the EvoTAR project was to develop generic and predictive models that allow a detailed description of the within-host dynamics and between-host dynamics of antibiotic resistance modules (genes, genetic elements, clones) and that will quantify the probability and rate of emergence and spread of resistance-conferring genes/mutations under various environmental conditions, different selective pressures and in different genetic backgrounds. To study the between host transmission of antibiotic resistance three generic model frameworks have been developed. With these models it is possible to study how different diseases and antibiotic resistances spread over a network of connected hosts (i.e. persons, hospitals, farms, etc.), to identify hosts at risk of becoming infected and for identifying hotspots for the emergence of multidrug resistant bacteria. Furthermore, results of the models could be used to aid in the development of nationwide surveillance programs by informing policy makers where to concentrate efforts. The use of *in silico* pharmacokinetic-pharmacodynamic (PKPD) models based on

data from *in vitro* time-kill experiments can provide valuable information to guide dosing of antibiotics. Experimental work on fitness costs has generated some general implications of importance for understanding and predicting resistance development. One important conclusion from this is that it is at present very difficult to predict the magnitude of the fitness effect of particular resistance mechanism in a particular genetic background.

Finally, novel interventional strategies to tackle antimicrobial resistance were developed and thoroughly evaluated. DAV132, developed partly in the context of EvoTAR by Da Volterra, is the first product with a clinically-demonstrated protection of intestinal microbiota from disruption during antibiotic treatments.

Apart from its scientific outcomes during the course of the EvoTAR projects important novel tools for the analysis of resistance genes, and the natural history of plasmids via a tool called PLACNET. A new method that we name pairwise comparative modelling (PCM) was developed to identify ARDs in large complex datasets. With ARD many novel resistance genes were identified and expanded very significantly the list of resistance known genes. Furthermore, the EvoTAR consortium successfully developed the experimental and computational workflows to use PacBio for reading out results of functional selections. This development enabled unprecedented quality and throughput of functional selections. A majority of the functional selections for this project rely on this approach. A novel *Targeted capture approaches* was developed in EvoTAR enabling cost-effective and high-throughput resistomes analysis. An antibiotic resistance gene-capture platform was designed that uses the SeqCap Ez technology of Roche NimbleGene. This platform consisted of 80,000 targets involving allelic forms of resistance genes and genes associated with the backbone of mobile genetic elements able to contribute to the spread of resistance. Furthermore, a core genome MLST (cgMLST) scheme was developed for *E. faecium* to standardize current intra-laboratory surveillance of this nosocomial pathogen. cgMLST transfers genome-wide single nucleotide polymorphism (SNP) diversity into a standardized and portable allele numbering system that is far less computationally intensive but with the resolution of SNP-based analysis of whole-genome sequencing (WGS) data.

# **Summary of Project context and Objectives**

This project addresses the problem of antibiotic resistance in bacteria. Antibiotics are one of the most apparent success stories of modern medicine and have saved the lives of countless people that suffered from bacterial infections. However, the use of antibiotics has also led to the emergence of antibiotic resistance in bacteria, which is a major threat to human health as therapeutic options for treating infections by antibiotic-resistant bacteria are increasingly limited. It is generally appreciated that the emergence of antibiotic resistance is a complex problem accelerated by the overuse of antibiotics. However, antibiotic resistance is a natural biological phenomenon with many facets that are still poorly understood: what are important reservoirs of antibiotic resistance? How do resistant and non-resistant bacteria interact in these reservoirs? Which conditions promote the evolution and transfer of resistance? Expanding our knowledge on these aspects will provide novel leads to combat the emergence of antibiotic resistance.

**The overall purpose of EvoTAR was to increase the understanding of the evolution and spread of antibiotic resistance in human pathogens.** More specifically, EvoTAR aimed to characterise the human reservoir of antibiotic resistance genes ("the resistome") by investigating the dynamics and evolution of the interaction between resistant and non-resistant bacteria from the human microbiome and the interrelations of the human resistome with environmental, animal and food reservoirs of resistance genes. Novel methods were developed and used to quantify resistance transfer under controlled conditions in gene exchange communities. Mathematical modelling have been applied to predict gene flow between different reservoirs and, consequently, to make a prognosis of future resistance trends. Novel *in vitro* and *in vivo* models of antibiotic resistance evolution and transfer allowed the study of the efficacy of novel intervention approaches aimed at reducing selection and spread of antibiotic resistance.

To reach its main objective, the multi-disciplinary EvoTAR consortium pursued the following five research themes:



To study objective 1 "*To elucidate the dynamics and the evolution of the interaction between resistant and non-resistant bacteria from the human microbiome*" the EvoTAR consortium performed the following three studies:

- A. Metagenomic sequencing of the human microbiome during and after administration of antibiotics (WP1)
- B. Population dynamics of resistant and susceptible enterococci and *Enterobacteriaceae* during and after administration of antibiotics (WP2)
- C. Experimental adaptive evolution of resistance genes (WP3, WP6).

Ad A. Metagenomic characterization has revealed changes in both the phylogenetic diversity and the total gene repertoire of the human microbiome during and after cessation of antibiotic treatment. Understanding both short-term and long-term effects of antibiotic treatment on the diversity of the human microbiome is essential because it sheds light on which organisms of the human microbiome are resistant to antibiotic exposure during treatment and thus can potentially transfer their resistance(s) to other bacteria. Persistent perturbations of the gut microbial communities have been associated to numerous chronic diseases. Possibly, antibiotic treatments might lead to such perturbations. Our study of the dynamics of the gut microbiome exposed to antibiotics provided for the first time large-scale information about the longer-term effects of antibiotic therapy. Furthermore, by comparing the sequencing data with databases of antibiotic resistance genes we were able to quantitatively determine which resistance genes are present in the human microbiome before, during and after antibiotic treatment

Ad B. For the functional studies on antibiotic resistant organisms genes enterococci and *Enterobacteriaceae* as Gram-positive and Gram-negative marker organisms, respectively, were selected for this project. Both groups of bacteria are commensals of the gastrointestinal tract but can cause lifethreatening infections in hospitalized patients. Enterococci and *Enterobacteriaceae* have been implicated in the transfer of important resistance mechanisms (for example vancomycin resistance in enterococci and Extended Spectrum β-Lactamases [ESBLs] in *Enterobacteriaceae*) between human and non-human reservoirs and to other pathogenic bacteria such as *Staphylococcus aureus* and *Pseudomonas aeruginosa*. Enterococci and *Enterobacteriaceae* can therefore be considered as paradigms for the functional study of the emergence and spread of antibiotic resistance. By determining the changes in the population of these bacteria during exposure to antibiotics throughout hospitalization, we were able to determine the dynamics of the emergence of antibiotic-resistance genes and clonal lineages associated with multi-drug resistance of enterococci and *Enterobacteriaceae*. This has given detailed information on the dynamics and interactions between resistant and non-resistant bacterial populations during and after antibiotic exposure.

Ad C. In the EvoTAR project, we have also studied evolution of resistance genes and the stability of resistance genes in bacterial hosts. This has provided important information on the risks that these resistance genes spread easily through a bacterial population, which would significantly increase the risk of widespread dissemination of these resistance determinants, or that these resistance genes will be inherently limited to a small range of hosts.

For objective 2 "*To characterize antibiotic resistance genes from the human microbiome and to elucidate the interactions of the human microbiome with environmental, animal and food reservoirs of resistance determinants*" the following three studies were performed:

- A. Functional metagenomics of antibiotic resistance genes in the human microbiome (WP3)
- B. Detection and quantification of antibiotic resistance genes in human and non-human reservoirs (WP4)
- C. Genomic characterization of antibiotic resistant bacteria from different reservoirs (WP2, WP5)

Ad A. The functional repertoire of antibiotic resistance genes in the human microbiome using functional metagenomic approaches was studied and revealed great dynamic. This has lead to an in-depth description of the effects that hospitalization and antibiotic therapy have on the repertoire of antibiotic resistance genes that are harboured by the bacteria from the human gut.

Ad B. Using a gene sequence capture technology we have detected and quantified the load of antibiotic resistance genes and the genes associated with their transfer in both human and non-human reservoirs in a high-throughput fashion. This revealed the extent by which resistance genes from the human reservoir are also found in non-human niches, indicating a possible transfer of antibiotic resistance genes between human and non-human reservoirs.

Ad C. By high-throughput genome sequencing of enterococci and *Enterobacteriaceae*, which are important vectors for the spread of antibiotic resistance (Livermore, 2009) and by high-throughput culturing to select antibiotic-resistant bacteria, followed by subsequent characterization by genome sequencing we have identified and characterized bacteria that play a central role in acquiring and transferring antibiotic resistance genes between bacteria and between different environmental reservoirs.

Objective 3 "*To determine the transfer potential of antibiotic resistance genes to human pathogens and to assess the contributions of the environment and the genetic elements which carry the antibiotic resistance genes on the efficiency of transfer*" included the following studies:

- A. Determination of factors affecting the bacterial host range of resistance plasmids (WP4, WP6)
- B. Contribution of environmental conditions to the efficiency of transfer of antibiotic resistance genes (WP6)
- C. Analysis of fitness costs incurred by resistance mutations and resistance plasmid carriage and genetic adaptation of resistance plasmids to novel bacterial hosts (WP6, WP7)

Ad A. A critical issue for the understanding of plasmids as disseminators of antibiotic resistance is to define the bacterial host-range of each resistance plasmid. Plasmid maintenance in ecosystems (for example, the human gut) depends on the stability of the resistance plasmid in the different bacterial hosts, which is a crucial factor in assuring spread of the plasmid by horizontal gene transfer. This aspect has been studied in EvoTAR.

Ad B. In the EvoTAR project the conjugation efficiency of plasmids and their subsequent capability to propagate in different bacterial hosts has been determined by a variety of methods. The environmental signals that trigger the induction of conjugation has been identified, which ead to insights into the ecological circumstances in which horizontal gene transfer takes place.

Ad C. Considerable attention has been given to determine the fitness costs that are incurred by resistance mutations and the carriage of resistance plasmids and to determine the efficiency by which different plasmids can adapt to their bacterial hosts.

Studies for Objective 4 "*To generate integrated mathematical models, using data on the evolution, transfer and spread of antibiotic resistance genes, which can be used to describe the flow of antibiotic resistance genes between different environments and bacterial hosts and to predict future resistance trends*" involved:

- A. Development of mathematical models that describe the probability and rate of resistance development taking into account several levels of modular trait interactions (WP7)
- B. Development of generic and predictive models which will lead to a detailed description of the within-host dynamics and between-host dynamics of antibiotic resistance modules and spread of antibiotic resistance at the population level (WP7)

Ad A. The success of an antibiotic resistant clone is largely determined by the competitive fitness of that clone in comparison to susceptible ones in a number of different environments. This overall fitness will be determined by many modular traits (the host bacterium, the plasmids, transposons or integrons with their resistance genes in the bacterial cell and the resistance gene itself) that act at different levels and which are studied as part of the other objectives in this proposal. EvoTAR studied how traits influence the survival, growth and transmission success of the various genetic elements that can influence the trait or are influenced by the trait using mathematical models that describe the probability and rate of resistance development taking into account several levels of modular trait interactions.

Ad B. A major objective of the EvoTAR project was to develop generic and predictive models which will lead to a detailed description of the within-host dynamics and between-host dynamics of antibiotic resistance modules (genes, genetic elements, clones) and which will quantify the probability and rate of emergence and spread of resistance-conferring genes/mutations under various environmental conditions, different selective pressures and in different genetic backgrounds. This modelling-based approach has proven essential for the prediction of the risk that a given antibiotic resistance gene may successfully spread to pathogens and thereby contribute to future resistance problems and how changes in the selective pressures influenced rates of spread of antibiotic resistance at the population level.

Finally studies indicated below were executed for Objective 5 "*To explore novel intervention approaches aimed at reducing the emergence and spread of antibiotic resistance*":

- A. Assessment of the efficacy of compounds that absorb and inhibit residual antibiotics in the colon to minimize emergence of antibiotic resistance (WP1, WP8).
- B. Assessment of the efficacy of compounds that impede the conjugative transfer of resistance genes among bacteria to minimize dissemination of antibiotic resistance (WP6, WP8)
- C. Identification of novel targets for therapeutic interventions (WP2)

Ad A. A novel intervention approach was conducted in EvoTAR aimed at administering a compound that absorb and inhibit residual antibiotics in the colon. It was anticipated and proven correct that this approach minimize selective pressures leading to the emergence of antibiotic resistance in the commensal flora without changing the fate of absorption of the antibiotic and its potential to treat the infection for which it has been administered.

Ad B. Horizontal gene transfer is widespread in the environment, where antibiotics are present at concentrations lower than those used in medicine, and where they perform other functions than those related to their therapeutic applications. The continuous presence of antibiotics and frequency of gene transfer make environmental microorganisms a good source for compounds capable of inhibiting transfer of genes associated with the response (including resistance) to different classes of antibiotics. In order to explore this possibility, a detailed characterization of natural and chemically synthesized compounds was tested for their capacity to inhibit conjugation.

Ad C. Using functional genomic approaches genes contributing to resistance that could serve as targets for the development of new therapeutic interventions were identified. The focus of these studies was oriented towards common gastro-intestinal commensals (enterococci and *Enterobacteriaceae*), that are a major hub for antibiotic resistance, by functional genomics-based approaches. The identification of the full complement of genes involved in antibiotic resistance in these groups of nosocomial pathogens will open up new avenues for the development of novel therapeutic interventions.

# **Main results**

The sensitivity of identifying antimicrobial resistance determinants (ARD) in large complex datasets like metagenomic datasets is low often resulting in identifying only a fraction of the genes that are actually present. Currently, assigning a protein a given function relies on the shared identity of the sequence (*i.e.* letters corresponding to amino acids) with a protein for which the function is known and certain (reference protein). Setting an identity threshold is problematic since many of the proteins identified in metagenomic datasets share a low identity with reference proteins. A low identity threshold would increase sensitivity yet leads to false positives while a high identity threshold would be specific but insensitive. Previous studies identified between 100 and 1093 ARDs in the Human intestinal microbiota by BLAST coupled with an identity threshold varying from 50% to 80% with a reference protein of the ARD family.

# **Census of ARDs of the human gut microbiome**

As opposed to primary-sequence similarities, comparison of the folded structures of proteins is expected to be more specific and sensitive. However, computing structures from primary sequence for a large number of proteins is challenging and was not carried out in a microbiome field. We developed in workpackage (WP)1 a new method that we name pairwise comparative modelling (PCM). PCM aims to predict protein functions with more specificity than one-dimensional method, and includes a large increase in sensitivity by allowing the functional assignment of proteins with low identity to known references. It relies on a highly efficient modelling algorithm, using an established ARD structure from the RCSB protein data bank as a positive template and the most closely related non-ARD structure as a negative template. We modelled the structure of all proteins encoded by a 3.9 million human gut microbial catalogue. A custom pipeline was developed to assess the fit of an unknown protein structure to a positive and a negative template. The approach was validated using a set of ARDs found in a search selecting antibiotic resistance from the soil microbiome; out of 1390 ARDs the pipeline correctly predicted 1374, a success rate of 96.6%. The average sequence identity was only 37.6%, well beyond that acceptable for BLAST-based searches.

The pipeline identified 6095 ARDs among the 3.9 million human gut microbial genes. They belong to various families. PCM outperformed other methods for all the families but the very well studied class A β-lactamases, where Resfams was more efficient.

Some 72% of ARDs could be assigned to a phylum; a majority were from *Firmicutes* (2962/6095, 48.6%%) and from *Bacteroidetes* (858/4405, 14.1%) while only 3.7% (225/6095) ARDs were from *Proteobacteria*,, that include most pathogens and were most studied previously. About 60% belonged to gene clusters denoted metagenomics units (that is, microbial genomes or sub-genome genetic elements); 95.6% of the clusters contain >500 genes and correspond to bacterial species, indicating that the majority of ARDs have chromosomal location.

In conclusion, identified a large number of novel ARDs and thus expanded very significantly the list of these important genes. Future studies will address the experimental validation of the ARD function.

With the establishment of an optimized approach to identify ARD we set out to study the dynamics of the resistome, *i.e.* the total of ARD, in the microbiota of patients upon hospitalization.

# **Impact of antibiotic treatment on the resistance determinants of the human gut microbiome**

We (WP1) analyzed the resistome upon two types of antibiotic exposures: patients suffering from cystic fibrosis (a chronic pulmonary disease that predisposes to bacterial infections and thus to a high, chronic antibiotic exposure) and patients receiving a selective digestive decontamination (SDD) antibiotic cocktail. As control, we analysed individuals studied in the MetaHIT consortium, which were not exposed recently to antibiotics. Among the control individuals (n=663) the abundance of the ARDs was significantly correlated with the overall gene and species richness of the microbiome. This fits the conclusion reached while establishing the list of the ARDs from the human gut, that most ARDs reside on bacterial species (see above) – the higher the number of species an individual harbours, the higher is the level of the ARDs he carries. The same relation between gut microbiome richness and ARD abundance was also found for the EVOTAR individuals not exposed to antibiotics (n=45) or present in a hospital ward but not treated with antibiotics (n=32).

A short-term exposure (SDD;  $n=10$ ) altered both the richness (a decrease of 67.8%) and the ARD abundance (a decrease of 65.2%) greatly. There was thus no enrichment of the ARDs, presumably because of the inability of most species to withstand the harsh antibiotic treatment (Figure 1).



Dynamics of richness and pARD richness under an intense and short exposure to antibiotics

Figure 1. Dynamics of gene richness and ARD richness under an intense and short exposure to antibiotics.

In contrast, a long term (chronic) exposure altered the relation between the richness of the microbiome, which was decreased and the abundance of the ARDs, which was increased. Clearly, such exposure selects for the species than can strive in constant antibiotic presence, due to the ARDs they encode.

In addition, to mapping ARDs from metagenomic datasets resistance genes in partly the same sample set were also identified using functional selections, work performed in WP3.

#### **Resistome mapping**

Our efforts to characterize the resistome in hospitalized patients and its dynamics using functional selections can be grouped into three distinct efforts: (1) To develop new techniques for more thorough and higher throughput characterization of resistance genes from gut microbiomes, (2) To catalogue the collection of resistance genes in the gut microbiome of hospitalized individuals, and (3) To assess the dynamics of the gut resistome during hospitalization and antibiotic treatment.

# **New techniques for more thorough and higher throughput characterization of resistance genes from gut microbiomes.**

#### *Parfums: High throughput metagenomic functional selections in E. coli*

During the course of the EvoTaR project new techniques have been developed that increase the throughput of metagenomic functional selection in *E. coli.* Our first development was to develop a experimental and computational workflow for utilizing the short reads that result from the Illumina sequencing platform. This approach developed together with the Dantas Lab is termed Parfums and was published in Science (Forsberg *et al* (2012) Science) along with the first application of the methodology to study the resistome of soil bacteria.

#### *PacBio based metagenomic functional selections in* E. coli*.*

Following the development of Parfums sequencing platforms continued to improve and we switched to use the PacBio sequencing platform as it provides long reads that span the entire insert of functional selection libraries. Working together with the sequencing facility at University of Oslo we successfully developed the experimental and computational workflows to use PacBio for reading out results of functional selections. This development enabled unprecedented quality and throughput of functional selections. A majority of the functional selections for this project rely on this approach.

#### *Metagenomic functional selections in* Lactococcus lactis

Most functional metagenomic studies deploy *E. coli* as the expression host due to its amiability for cloning. Using a Gram-negative host strain is likely to introduce bias in the specific genes that are identified in functional metagenomic selections. To assess the extent to which this bias is causing significant problems we developed an experimental workflow for using *L. lactis* as our cloning host. Notably, this required substantial optimization to achieve sufficient library sizes. However, the results of these experiment remained discouraging as no resistance determinants other than those found in an *E. coli* small-insert library (specifically against the antibiotics tetracycline and D-cycloserine) could be detected in the *L. lactis* metagenomic libraries. Importantly, resistance determinants against the antibiotics linezolid, vancomycin and daptomycin (which are active against Gram-positive bacteria, but not against Gram-negatives) have not been identified. We conclude that there is currently no added value in further developing functional metagenomic libraries in *L. lactis* and have consequently terminated this particular research line.

#### **Catalogue the collection of resistance genes in the gut microbiome of hospitalized individuals**

#### *Cataloging resistomes from* E.coli *small insert functional selections*

Metagenomic expression libraries from over 60 patient samples have been constructed in the Gramnegative host *E. coli*. Screening for resistant clones from these libraries have resulted in the identification of tens of thousands different clones. All metagenomic inserts from resistant clones have been sequenced and annotated using the new PacBio based approach described above. In addition a novel annotation pipeline has been created to rapidly analyse metagenomic insert sequences derived from PacBio sequencing. A catalogue of resistance genes from the ICU patients has been constructed on the basis of these efforts. This catalogue contains over three thousand genes. Interestingly, a majority of these resistance genes are closely related to previously identified resistance genes. Analysis of the context of these resistance genes revealed that over 30 % of them have been previously identified within human clinical isolates.

#### *Cataloging resistomes from* E. coli *fosmid functional selections*

Using fosmid libraries we have identified genes that confer resistance to the disinfectant benzalkonium chloride (BC) from the human gut microbiota. Two of the genes that conferred BC-resistance to *E. coli* were predicted to be involved in membrane transport or efflux and to originate from Gammaproteobacteria and Bifidobacterium respectively, whereas the third gene was predicted to function

as an UDP-glucose-4-epimerase, originating from *Eggerthella lenta*. Two BC-resistant clones exhibited reduced susceptibility towards the antibiotics erythromycin and tobramycin, with one of these clones also showing reduced susceptibility to ampicillin. These data show that the human gut microbiota is a reservoir for genes that confer resistance to disinfectants. The reduced susceptibility to antibiotics in two BC-resistant clones indicates that, in gut bacteria, resistance to BC can be genetically linked to resistance against antibiotics.

# **The dynamics of the gut resistome during hospitalization and antibiotic treatment**

# *Resistome dynamics from fosmid libraries and real time PCR*

A preliminary analysis of the resistome dynamics was conducted for a subset of the patient samples for which samples existed from the first sampling point at the hospital, during hospitalization and after discharge. The results revealed that the resistome was highly dynamic and expanded during hospitalization in terms of abundance (e.g. more antibiotic resistant clones where selected from the libraries). These results are published in Journal of Antimicrobial Chemotherapy (Bulow *et al* (2014), JAC).

# *Resistome dynamics of key resistance genes based on microfluidic real time PCR*

We (WP3) set up a high-throughput nanolitre-scale real-time PCR assay (using the 96.96 BioMark™ Dynamic Array for Real-Time PC, developed by Fluidigm Corporation, San Francisco, CA, U.S.A) to detect and quantify the presence of 85 resistance determinants in metagenomic DNA. We used this platform to characterize the dynamics of the resistome of the gut microbiota of patients during hospitalization and to assess the spread of resistance genes through sewage. This technique allows highthroughput characterization of key constituents of the gut resistome during antibiotic treatment and provides a low cost alternative to using functional metagenomics to characterize resistome dynamics.

# *Resistome dynamics based on comprehensive small insert libraries*

Finally, we analyzed the dynamics of the gut resistome on the basis of the small insert functional selection libraries described above. This analysis revealed that the gut resistome is subject to substantial dynamics during antibiotic exposure and hospitalization. Two main conclusions where drawn from this analysis:

- (i) The gut resistome expands during hospitalization, yet recovers after discharge. The abundance and diversity of antibiotic resistance genes identified in functional selection experiments more than double from the first sampling point (admission to ICU) to the subsequent sampling points during hospitalization. Then following approximately 6 months after discharge, the abundance of antibiotic resistance genes return to the same level.
- (ii) The gut resistome becomes more enriched in resistance genes that are also shared with human pathogens during hospitalization. This is consistent with the hypothesis that the gut resistome acquires pathogenic resistance genes during hospitalization.

In conclusion we can see that hospitalization and antibiotic treatment has profound effects on the gut resistome. Furthermore, the expansion of the resistome during hospitalization could lead to an increased risk of transfer of antibiotic resistance genes to infecting pathogens.

In another culture independent method to analyse the resistome of human reservoirs (hospital and community) and reservoirs with a direct link to humans (foodborne animals) we (WP4) developed and used a next-generation *one-step* high-throughput targeted platform that was designed and validated during the EvoTAR project. Analysis of *resistome* uses metagenomic approaches, either "open" (target gene sequencing, shotgun metagenomic sequencing, metatranscriptomic sequencing) or "closed" formats (targeted and/or functional gene arrays) show poor sensitivity and specificity, and limited quantitation possibilities. *Targeted capture approaches* are the more cost-effective and high-throughput alternatives to obtain large data sets of orthologous genes from many individuals and was chosen to enhanced the resistomes analysis in the EvoTAR project. Advantages of targeted platforms, especially the new generation of *in-solution* targeted capture platforms, over array-based platforms or other genome-

partitioning are scalability, cost-effectiveness, and enhanced data quality (lower variance in target coverage, more accurate SNP calling, higher reproducibility and longer assembled contigs). Furthermore, current resistome analyses do not usually take into consideration either the genetic elements involved in gene mobility or the genes that contribute to co-selection of AR as those encoding resistance against biocides or heavy metals. We developed an AR gene-capture tool based on current knowledge on the genetic structure of resistance genes and the platforms to be propagated (about 80,000 allelic forms, including genes associated with the backbone of mobile genetic elements able to contribute to the spread of resistance), with the aim of reaching the desirable level of comprehensivity and curation required to analyze in depth the resistomes of different ecosystems. The platform uses the SeqCap Ez technology of Roche NimbleGene and involves both academic and industry collaborations. The company has expressed interest in the commercialization of this custom platform. We conclude that WP4 provided significant scientific and technological outcomes.

#### **Development of a custom targeted capture platform (TCP) to analyze resistomes.**

A customized SeqCap EZ platform, a solution-based capture system that allows the enrichment of genes or genomic regions in a single test tube, was designed to capture the resistome and mobilome of fecal microbiomes of humans (hospitalized and non-hospitalized) and foodborne animals. This large subtask involves different steps: i) analysis of current databases, ii) design of a custom SeqCap EZ platform, iii) validation of the platform. These activities were fully accomplished in 2014 and further optimized in 2015.

#### **Analysis of DBs and creation of a curated DB.**

A deep analysis of DBs for antibiotic resistance (CARD, http://arpcard.mcmaster.ca/; ARGannot http://en.mediterranee-infection.com/article.php?laref=283&titre=arg-annot-), antimicrobial biocide and heavy metals (BACMET, http://bacmet.biomedicine.gu.se/), and relaxases was performed to build a homemade non-redundant database of well-known genes. The workflow consisted on building a homology network (Blast All-to-All of protein sequences), and performing a cluster analysis and further manual annotation of AbR families. All proteins of each cluster were aligned to obtain profile Hidden Markov Models (HMMs) for each family. Then, two phylogenetic analyses were done. One included the proteins of each AbR family, and the other included proteins of each AbR family and their associated profile HMMs against Uniref100 database (http://www.uniprot.org). The last step is essential to improve annotation and identify both false positives and potential undescribed genes.

# **Design and manufacturing of the platform.**

The number of targets included in the platform was significantly increased in comparison with the initial contrive described in the DoW (from 1,000 to 81,000 targets) taking advantage of the development of one-step targeted capture methods in the last three years, the bioinformatic tools and the improvement in the gene DBs (curated DBs used as template for our design became available after 2012). The final version of the SeqCap EZ capture platform consists of 7,963 non-redundant AbR genes, 47,806 manually curated genes from the results of profiles HMMs against Uniref100; 30,740 genes from BacMet-Antibacterial Biocide & Metal Resistance database, and 2,517 non-redundant genes from the home-made relaxase database provided by partner 10 (UC). All genes assemble a platform capture of approximately 81,000 non-redundant genes. This SeqCap Ez custom platform for capturing AbR, Metal and MGE has been manufactured in collaboration with RocheNimbleGen Inc (Madison, USA; www.nimblegen.com). A full bioinformatics workflow was developed as a complement of the design of the SeqCap EZ platform (Data processing and Functional annotation). First, the reads from the sequencing are filtered using the capture platform as a reference to remove the rubbish sequences. Then the filtered reads are assembled to make contigs. Finally, we use GeneMark to ORF prediction. Functional annotation consists of three steps. Annotation against capture platform, selected genes are double annotated against general database (i.e. Uniprot or RefSeq). If the candidate genes still pass the double filter as resistant gene (AbR, biocide or Metal) or as relaxase, the gene is studied phylogenetically to establish if this gene are a novel resistance gene or a known gene.

# **Validation of the platform.**

A pilot to estimate the coverage of genes included in the final version of the customized SeqCap EZ platform, the suitability of different protocols for library preparation (Kappa vs Nextera approaches), NGS sequencing (NextSeq 2x100 changing the number of samples per run) and also the bioinformatic analysis of metagenome.

On-target capture ranges from 30 to 50% of the sequencing reads in the samples analyzed to date, data being within the range obtained for other SeqCap EZ platforms used in human genetics. In comparison with traditional metagenomics methods, the EvoTAR platform improves *sensitivity* (250-fold increase), *specificity* and detection of gene *diversity* (see Figure 2, 3, 4, and 5).



Figure 2. Box plot showing the number of detected genes per million of reads using the EvoTAR TCP of the samples form animals and hospitalized analysed. The two panels represent the number of gene alleles coding for resistance to antibiotics, which are plotted as box plots.



Fig. 3. Dot-plot of the genes-based alignment corresponding to the number of reads captured by conventional metagenomics (y-axis) and to the number of reads captured in the experiments with the current EvoTAR platform prototype using SeqCap EZ Illumina technology (values x-axis). Reads corresponding to the same gene sequence are aligned and represented by dots. Each color represents different metagenomes analysed (in this figure represented by samples from swine). An average of 15.0 million 100 bp paird reads were obtained for each individual sample in the SeqCap EZ Illumina experiments.



Figure 4. Box plots showing the differences in specificity and sensitivity of gene detection using conventional metagenomics (pre-capture) versus the EvoTAR TCP (post-capture), in this figure represented by samples from swine. The three panels represent the number of gene alleles coding for resistance against antibiotics, metals and biocides, or genes associated with mobile genetic elements (relaxases), respectively, which are plotted as box plots.



Figure 5. Diversity of the genes conferring resistance to antimicrobials using shotgun metagenomics vs targeted capture platform

## **A comprehensive outline of the AbR genes in animals and humans.**

Samples from humans and animals contain genes coding for resistance to nine antibiotic families, namely beta-lactams (33 groups), aminoglycosides (33 groups), macrolides (23 groups), tetracyclines (23 groups), fluoroquinolones (5 groups), sulfonamides (3 groups), trimethoprim (11 groups), glycopeptides (5 groups), and chloramphenicol (13 groups). For the first time, we can identify with high sensitivity a wide diversity of resistance genes using metagenomic samples.

The diversity of AbR genes found in animals (consider that this set only represents a small sample) is higher than that found in samples from hospitalized patients (Fig. 2), with remarkable diversity of genes conferring resistance to tetracycline. Many AbR were found in both hosts but others, which include genes of all antibiotic families, were more abundant or associated with any of them.

#### **Dynamics of the resistome.**

Changes in the composition of AbR genes in samples form hospitalized patients were observed, with remarkable gains and loss in the recovery of certain families of genes.

Next to do culture-independent methods described above work performed in the framework of WP5, concentrated on developing and applying innovative high-throughput culturing strategies, including the MicroDish platform, for the isolation of antibiotic resistant microbial populations from the human intestinal tract as well as from other gut and non-gut environments. In addition, tailored narrow-spectrum high throughput cultivation has been employed to target microorganisms with specific phenotypes of interest, such as resistance to vancomycin and carbapenem.

#### **Optimized cultivation methods for the human gut microbiota**

Optimized cultivation methods for the human gut microbiota were developed that can capture a representative majority of the cells present in a sample by both abundance and overall community structure. To this end, a new method was established to perform profiling of antibiotic resistance of the gut microbiota in a high-throughput multiplex fashion allowing simultaneous profiling of 16 types of antibiotic resistance in gut bacteria. These resistance profiles provided the means to target bacteria identified as of great interest by the Human Microbiome Project (HMP) and culture several previously uncultivated bacteria. The genomic analysis of several strains from the Human Microbiome Projects most wanted list confirmed the multidrug resistance profiles of the strains as well as the novelty of the strains in relation to previously sequenced organisms. The organisms are currently being characterized to ensure proper phylogenetic classification including naming of new species.

# **High throughput cultivation screens to retrieve antibiotic resistance bacteria from humans, animal and the environment**

High throughput cultivation screens have been used to retrieve antibiotic-resistance bacteria from human and animal intestinal tracts as well as from marine environments known for their prolific production of

antibiotics, and thus can be expected to also represent natural hubs for antimicrobial resistance. Previously uncultured bacteria could be cultivated using conventional plating, and application of antibiotics in the media can serve to capture a greater bacterial diversity. Moreover, we developed criteria to address an important caveat of the plate scraping method whereby bacteria may be detected that did not actually grow. Furthermore, genomic DNA was isolated from the regrown isolates in order to allow for a functional screen for resistance genes using small insert library screening. The cultivation study in which antibiotic-resistant bacteria were isolated from sponge samples yielded >200 strains that are *Pseudovibrio* spp. There is special focus on *Pseudovibrio* isolates since they are multidrug resistant, have the potential to produce antimicrobial compounds, and were isolated in large numbers from the different sources. Based on GTG-5 genomic fingerprinting and resistance profiles, 25 different strains were selected for whole-genome sequencing. All four type strains (*Pseudovibrio ascidiaceicola* DSM 16392, *Pseudovibrio axinellae* DSM 24994, *Pseudovibrio denitrificans* JCM 12308 and *Pseudovibrio japonicus* NCIMB 14279) belonging to this this genus were also sent for whole-genome sequencing.

A second cultivation-based study using the MicroDish platform was done wherein anaerobic antibioticresistant bacteria were isolated from faecal samples (n=20) obtained from SDD patients received from Partner 1, with the aim to broaden our knowledge regarding the role of little-studied and/or novel anaerobic bacteria as a reservoir for antibiotic resistance.

Furthermore, we examined the potential for antibiotic production by assessing the expression of associated secondary metabolite biosynthesis gene clusters. Metatranscriptome datasets from intestinal microbiota of four human adults, one human infant, 15 mice and six pigs, of which only the latter have received antibiotics prior to the study, as well as from sea bacterioplankton, a marine sponge, forest soil and sub-seafloor sediment, were investigated. We found that resistance genes are expressed in all studied ecological niches, albeit with niche-specific differences in relative expression levels and diversity of transcripts. For example, in mice and human infant microbiota predominantly tetracycline resistance genes were expressed while in human adult microbiota the spectrum of expressed genes was more diverse, and also included β-lactam, aminoglycoside and macrolide resistance genes. Resistance gene expression could result from the presence of natural antibiotics in the environment, although we could not link it to expression of corresponding secondary metabolites biosynthesis clusters. Alternatively, resistance gene expression could be constitutive, or these genes serve alternative roles besides antibiotic resistance.

# **Role of soil microbiota in the origin and evolution of resistance**

Investigations on the role of the soil microbiota in the origin and evolution of resistance to two critically important antimicrobial classes in human medicine, glycopeptides and carbapenems led to i) description of a new glycopeptide resistance operon in *Rhodococcus equi* and ii) discovery of multiple new carbapenem-hydrolyzing enzymes produced by environmental bacteria.

Glycopeptides such as vancomycin and teicoplanin are last resort drugs for treatment of nosocomial infections caused by methicillin-resistant *Staphylococcus aureus* and enterococci. Although progenitors of vancomycin resistance genes have been described in soil bacteria, elucidation of the evolutionary trajectories leading to the presence of vancomycin resistance genes in clinical isolates is still lacking. We characterized the vancomycin resistance operon of a *Rhodococcus equi* isolated from soil in Denmark in 2004 and displaying a vancomycin-resistance phenotype inducible by glycopeptides. The vancomycin resistance operon in this strain had an unique genetic organization and was designated as *vanO.* This operon has low homology to enterococcal *van* operons and harbors a *vanHOX* cluster transcribed in opposite direction to the *vanS-vanR* regulatory system and comprised between three open reading frames with unknown function. This finding has clinical interest since glycopeptides are used to treat *R*. *equi* infections and resistance has been reported in clinical isolates.

Furthermore, detection and functional characterization of carbapenem-hydrolyzing ß-lactamase in environmental bacteria was achieved by antibiotic selective culture, functional metagenomics and sequence database mining. The culture-based approach yielded 29 bacterial isolates from 13 soil samples. Among these, we detected isolates belonging to genera or species for which MBL production was not reported prior to this study. Seven new metallo-ß-lactamases (MBLs) were discovered in *Pedobacter* 

*roseus* (in which the produced CHß was annotated as PEDO-1), *Pedobacter borealis* (PEDO-2), *Pedobacter kyungheensis* (PEDO-3), *Chryseobacterium piscium* (CSP-1), *Epilithonimonas tenax* (ESP-1), *Massilia oculi* (MSI-1), *Sphingomonas* sp. (SPG-1) and *Epilithonimonas tenax* (ESP-1). Plasmid libraries were constructed from 10 of the soil samples used for the culture approach. We detected two subclass B1 MBLs (annotated as  $bla_{GRD23-1}$  and  $bla_{SPN79-1}$ ) and 7 subclass B3 MBLs ( $bla_{CRD3-1}$ ,  $bla_{OSNS-1}$ ,  $bla_{\text{GRD33-1}}$   $bla_{\text{OSN49-1}}$ ,  $bla_{\text{ALG6-1}}$ ,  $bla_{\text{ALG11-1}}$ , and  $bla_{\text{DHT1-1}}$ ) in six of the 10 soil samples analyzed. The crude extract of the metagenomic MBLs showed significant imipenem hydrolysis. Taxonomic classification at the phylum level by RAIphy suggested that six enzymes originated from Proteobacteria, two from Bacteroidetes and one from Gemmatimonadetes. The sequence database mining approach allowed identification of three new resident carbapenemases in *Chromobacterium* sp*.* strain C-61 (CRS-1), *Chromobacterium haemolyticum* DSM 19808 (CRH-1) and *Chromobacterium piscinae* ND17 (CRP-1), which are species commonly found in aquatic environments. These carbapenemases showed between 68 and 76 % amino acid identity to *Klebsiella pneumoniae* carbapenemase (KPC), suggesting that *bla*<sub>KPC</sub> may have evolved from possible ancestor genes resident on the chromosome of members of the genus *Chromobacterium*. Overall we isolated 17 new carbapenemases with 30-76% amino acid identity to previously confirmed carbapenemases. Our results show that the environment including soil and water is a source of highly diverse carbapenemases that are produced by a variety of bacterial species and have not yet emerged in clinical settings. These carbapenemases may constitute potential carbapenemresistance determinants of clinical relevance if acquired by pathogenic bacteria, as they were functionally expressed in *E. coli*.

#### **Culturing antibotic resistant isolates from the environment**

Two hundred strains were cultured from marine environments. Phenotypic analysis suggests that some of the marine isolates (generally epiphytes) are exceptionally broadly resistant to antibiotics (both intrinsic and determined by individual genes) including beta-lactams, colistin (the basis of selection) but also chloramphenicol, kanamycin and derivatives and many others. Further, whilst sensitive to both rifampicin and erythromycin spontaneous resistance emerges at high frequency and in some cases (especially rifampicin resistance) correlated with changes in other phenotypes related to cell organization. One such strain, a *Flavobacterium*, is being focused on in depth. Additionally, resistance to non-clinical antimicrobials, accumulating resistant marine strains, was studied. A strategy of cloning functional resistance genes into *E. coli*, both from beta-lactam resistant strains and for antimicrobials of unknown mechanism (derived from our collection of marine bacteria) is being pursued with success in isolating resistant strains. Library construction is in progress from 10 strains. This effort uses co-culture using MDCC with producer and target strains, often working with limited quantities of antimicrobials. The aim is to identify resistance genes of unusual antimicrobials (antibiotic candidates) and provide insight into mechanism of action. The strain collection now stands at >250 strains. Finally, the ability of swarming bacteria to contribute to the spread of antibiotic resistance was studied. To this end, Swarms of the flagellated bacterium *Paenibacillus vortex* have been shown to collectively transport other microorganisms. It was found that *P. vortex* can invade antibiotic-rich environments by carrying antibiotic-degrading bacteria; this transport is mediated by a specialized, phenotypic subpopulation utilizing a process not dependent on cargo motility. Swarms of beta-lactam antibiotic (BLA)-sensitive *P. vortex* used beta-lactamase-producing, resistant, cargo bacteria to degrade BLAs in their path. In the presence of BLAs, both transporter and cargo bacteria gained from this temporary cooperation; there was a positive correlation between BLA resistance and dispersal. *P. vortex* transported only the most beneficial antibiotic-resistant cargo (including environmental and clinical isolates) in a sustained way.

In addition to the studies described above that aimed at identifying resistant organisms and/or resistance genes in complex ecological entities like the human gut or soil samples, the work performed in WP2 focused on understanding the evolution of Enterobacteriaceae and enterococci from commensal organisms to multidrug-resistant opportunistic pathogens. Bacteria can become resistant to antibiotics by the acquisition of resistance genes or by the accumulation of mutations in the target of an antibiotic, but in order to become successful nosocomial pathogens additional adaptations are often required, e.g. those that contribute to the efficient colonization (and infection) of hospitalized patients.

In this WP EvoTAR scientists worked to understand the evolutionary trajectories that contribute to the emergence of successful clones of multi-drug resistant nosocomial pathogens, using both comparative and functional genomic approaches. In comparative genomic studies, high-throughput sequencing approaches are used to sequence the genomes of a number of strains, in order to determine their evolutionary relatedness and their repertoire of antibiotic resistance genes, mutations associated with antibiotic resistance and other adaptive elements. In functional genomic studies, high-throughput approaches (e.g. transposon mutagenesis and RNA-seq) are used to efficiently study the function of genes in a bacterium. The work in WP2 was divided in three different tasks and the results of these will be summarized below.

## **Comparative phylogenomics**

Work under this task focused on the multi-drug resistant nosocomial pathogens *Escherichia coli*, *Klebsiella pneumoniae* and *Enterococcus faecium*.

#### **Comparative genomics of ESBL-***E. coli***.**

A total of 24 ESBL-positive *E. coli* strains from chickens, chicken meat and humans and eight ESBLpositive *E. coli* strains that were isolated from pigs and from farmers that tended to these pigs, were sequenced and analyzed to assess whether resistant strains and/or antibiotic resistance plasmids are able to spread between these different niches. The results of these analyses revealed that genome sequencing could differentiate strains that are indistinguishable by classical sequence-based typing methods. Indeed, the differences within strains that were considered to be identical (i.e. had the same sequence type (ST)), were substantial (more than 4200 SNPs that could be identified), indicating considerable evolutionary divergence between strains in a single ST, and low-likelihood of recent transmission between poultry and humans. This observation is important as this contradicts statements in previous studies, which have claimed that identical strains were spreading from chickens to humans. Additional analyses of the genome sequence data, which were performed in collaboration with the group of Prof. Fernando de la Cruz (University of Cantabria, partner #10 in EvoTAR), have focused on the sequence-based reconstruction of plasmids that carry ESBL genes. This approach revealed that virtually identical ESBLcarrying IncI1/ST3 and IncI1/ST7, as well as AmpC-type β-lactamase-carrying IncK plasmid backbones were shared by genetically unrelated human and poultry isolates, strongly suggesting that ESBLs are mainly disseminated via epidemic plasmids that can spread between different reservoirs.

#### **Comparative genomics of** *E. faecium*

In a collaboration with groups at Harvard Medical School (Boston, USA) and the Broad Institute (Cambridge, USA), 51 *E. faecium* genomes were sequenced and analysed. This study revealed that the *E. faecium* population can be split into two major clades (termed A and B), in line with previous analyses. However, our data showed that clade A can be further split in two sub-clades (A1 and A2). Interestingly, the large majority of modern clinical isolates can be assigned to clade A1, while strains from foodproducing animals are part of clade A2. Both strains from clade A1 and A2 can acquire resistance to vancomycin, while strains from clade B, which are typically isolated from healthy humans, are not resistant to vancomycin. Additional work on *E. faecium* specifically identified recent recombination events with the adaptation of *E. faecium* to novel niches, such as antibiotic-treated hospitalized patients. Interestingly, flows of recombination were largely congruent with the previously determined genetic subpopulations of *E. faecium*. Particularly, genes that are potentially involved in virulence (such as those coding for the biosynthetic machinery for capsule) or antibiotic resistance, specifically a D-alanyl-Dalanine carboxypeptidase that contributes to resistance to β-lactam antibiotics, are located in genomic regions that are of recombinogenic origin.

In another comparative genomic study, whole-genome sequencing was used to characterize how clinical *Enterococcus faecium* strains evolve during long-term patient gut colonization. To this end, the genomes of 96 *E. faecium* gut isolates, obtained over 8 years from 5 different patients, were sequenced. In addition to these 96 genomes, we also included publicly available genome sequences of 70 *E. faecium* strains to comprehensively describe *E. faecium* genome dynamics. All of the 96 patient isolates were grouped in *E. faecium* clade A, with only one strain clustering in clade A2. The remaining 95 strains were assigned to clade A1. The phylogenetic tree showed 5 clusters of closely related strains of patients, revealing the

microevolution of *E. faecium* strains during gut colonization. Evidence for direct transfer of strains between patients during hospitalization in the same ward was also obtained. In addition to core-genome based analyses, gene gain and loss was also studied, showing that loss and gain of prophages and plasmids is an important factor in generating genetic diversity during gut colonization. This study highlights the ability of *E. faecium* clones to rapidly diversify, which may contribute to the ability of this bacterium to efficiently colonize new environments and rapidly acquire antibiotic resistance determinants.

A core genome MLST (cgMLST) scheme was developed for *E. faecium* to standardize current intralaboratory surveillance of this nosocomial pathogen. cgMLST transfers genome-wide single nucleotide polymorphism (SNP) diversity into a standardized and portable allele numbering system that is far less computationally intensive than SNP-based analysis of whole-genome sequencing (WGS) data. The *E. faecium* cgMLST scheme was built using 40 genome sequences that represented the diversity of the species. The scheme contained 1.423 cgMLST target genes and was tested using WGS analysis of 103 outbreak isolates from five different hospitals in The Netherlands, Denmark and Germany. The cgMLST scheme performed well in distinguishing between epidemiologically related and unrelated isolates, even between those that had the same 'traditional' sequence type, which denoted the higher discriminatory power of this cgMLST scheme over conventional MLST. The *E. faecium* cgMLST scheme's performance was found to be equivalent to a SNP-based approach. The cgMLST scheme will facilitate rapid, standardized, high-resolution tracing of *E. faecium* outbreaks.

#### **Comparative genomics of** *Klebsiella pneumoniae***.**

Genome sequencing was performed of three pairs of KPC-producing *K. pneumoniae* strains which were isolated from patients. KPC stands for *K pneumoniae* carbapenemase; this enzyme confers reduced susceptibility to resistance to all β-lactam antibiotics including penicillins, cephalosporins, and carbapenems. The pairs of KPC-producing *K. pneumoniae* strains consisted of a colistin-susceptible strain that was isolated from a patient prior to colistin therapy and a colistin-resistant strain from the same patient that acquired colistin resistance during therapy. In colistin-resistant isolates a non-synonymous mutation in the *pmrB* gene and a disruption of the *mrgB* gene, which was proposed to be involved in the regulation of the PhoP/PhoQ two component system, were identified. These genome-based findings were corroborated by the construction of targeted mutants in *K. pneumoniae*, indicating that the upregulation of the PhoQ/PhoP system and activation of the *pmrHFIJKLM* operon which leads to resistance to polymyxins in *K. pneumoniae* through modification of lipopolysaccharides. The genome sequence data were used to assess how widespread the identified colistin resistance mechanisms are in a collection of 55 colistin-resistant clinical strains of KPC-producing *K. pneumoniae* from Italy and Greece. This analysis showed that *mgrB* inactivation is a common mechanism (detected in approximately 50% of cases), which highlights its clinical relevance. Moreover, the study also showed that different mechanisms of *mgrB* inactivation can be responsible for clinical resistance.

#### **Species-selective metagenomic analysis**

In this task we aimed to study the dynamics of the *Enterococcus* and *Enterobacteriaceae* sub-populations in the gut microbiota during hospitalization. Because metagenomic shotgun sequencing cannot accurately detect shifts in species that are present in relatively low levels in feces, culture enrichment needs to be performed. To determine which culture media were appropriate for this study, three enrichment media (Kanamycin Aesculin Azide broth, Enterococcosel broth and Enterobacteriaceae Enrichment broth) and a non-selective medium (Brain Heart Infusion broth) were used to enrich *Enterococcus* and *Enterobacteriaceae* sub-populations from fecal samples of healthy human donors. Enrichment was determined using multiplexed, high-throughput 16S rRNA sequencing on the Illumina MiSeq platform. While in all samples *Enterococcus* and *Enterobaceriaceae* were present at <1% of the population, enrichment led to >60% of *Enterobacteriaceae* in all four samples and >60% of *Enterococcus* in three out of four samples (enrichment for enterococci was unsuccessful in the fourth sample). Based on these data, Enterococcosel broth and Enterobacteriaceae Enrichment broth were chosen for further experiments. Enrichment cultures were performed with fecal samples from seven patients during hospitalization at the Intensive Care Unit. While enrichment for *Enterobacteriaceae* proved to be largely unsuccessful in this patient population (possibly because *Enterobacteriaceae* are eradicated from the gut

microbiota of these ICU patients due to prophylactic antibiotic therapy), enterococci could be enriched from nine out of sixteen samples. DNA was isolated from the fecal samples and these samples have been sequenced. However, analysis has been delayed because non-enriched samples from hospitalized patients and a trial in WP8 were prioritized. Data analysis is taking place currently.

## **Functional genomics of enterococci and** *Enterobacteriaceae*

A system for the generation of a library of transposon resistance mutants in *E. faecium* was developed and initially coupled to a microarray-based transposon mapping approach. This approach led to the identification of three novel determinants of ampicillin resistance in *E. faecium*. The identified intrinsic ampicillin resistance genes are highly conserved among *E. faecium* strains, indicating that this organism has a high potential to evolve towards ampicillin resistance. Using the same high-throughput functional genomic screening, a two-component system was identified in *E. faecium*, which contributes to decreased susceptibility to the disinfectant chlorhexidine and the antibiotic bacitracin. Subsequently, a method for next-generation sequencing based screening of transposon mutant libraries (Tn-seq) was developed for the vancomycin-resistant isolate *E. faecium* E745. This method was used to identify genes that are important for vancomycin-resistance in *E. faecium*. As part of this project, long-read PacBio and Oxford Nanopore sequencing was used to complete the draft genome sequence of *E. faecium* E745, which was previously sequenced by short-read sequencing (Illumina). Several methods for the computational analysis of Tn-seq data were developed (including the quantification of transposon insertions in a 25-nt window and a gene-per-gene analysis). Practically all genes that were identified to be contributing to vancomycin resistance in *E. faecium* E745 belonged to the vancomycin resistance transposon that is present in this strain. A small number of additional genes, with borderline statistical significance, were found to be putatively involved in vancomycin resistance in *E. faecium* E745. However, the targeted deletion mutants that were generated in these genes did not exhibit an increased susceptibility towards vancomycin, compared to the wild-type strains. These data suggest that the vancomycin resistance transposon is solely responsible for vancomycin resistance in *E. faecium*. It can, however, not be excluded that genes that are essential to *E. faecium* (and which therefore cannot be disrupted by a transposon) may functionally contribute to vancomycin resistance.

Changes in biocide and antibiotic susceptibilities, metabolism, and fitness costs were studied in biocideselected *E. coli* and *K. pneumoniae* mutants. Some strains that developed resistance to the disinfectant triclosan showed marked increases in MICs to several antibiotics, including ampicillin, ciprofloxacin and tetracycline. However, these mutants exhibited significant fitness costs in acquiring resistance. These various phenotypes suggest a trade-off of different selective processes shaping the evolution toward antibiotic/biocide resistance and influencing other adaptive traits.

A comprehensive transposon mutagenesis profiling was performed to identify genes that contribute to antibiotic resistance in the nosocomial pathogen *K. pneumoniae*. Effects on resistance (either increased susceptibility or increased resistance) have now been described and experimentally confirmed for 101 genes. Three mutants were found to exhibit increased susceptibility to cephalosporins and carbapenems. This result is relevant as *K. pneumoniae* is becoming increasingly resistant to these classes of antibiotics and it is of importance to understand the mechanisms by which resistance can emerge with the long-term goal to develop novel anti-infectives targeted against *K. pneumoniae.*

Expanding on the genomic analyses on colistin-resistant *K. pneumoniae* described above, loss-of-function mutations in *mgrB* were further characterized. These mutations are stable and occur without major consequences on fitness and virulence, leading to the efficient dispersal of this particular multidrugresistant *K. pneumoniae* clone. Interestingly, *K. pneumoniae* ST512 was found to have a higher mutation frequency than other clones, potentially contributing to its ability to become colistin-resistant.

# **Evolution of antibiotic resistance and fitness costs**

Studies in *Salmonella* were performed in WP2 to identify mutations that lead to antibiotic resistance (against streptomycin, colistin and meropenem) at concentrations 10-fold above the Minimum Inhibitory Concentration (MIC) and 0.2-fold below the MIC. Completely different mutations accumulate at low concentrations, compared to those that occur at high antibiotic concentrations. Nevertheless, even accumulated mutations at concentrations below the MIC lead to high-level resistance, indicating that exposure to low levels of antibiotics may still lead to the emergence of highly resistant bacterial populations.

In another study, the compensation of fitness costs of resistance (severely impaired growth rate associated with resistance due to absence of two major outer membrane porins) was studied in *E. coli*. An evolution experiment was performed with 16 lineages of *Escherichia coli in which the ompCF* genes were deleted with reduced fitness and increased resistance to different classes of antibiotics, including the carbapenems ertapenem and meropenem. After serial passaging, the relative growth rate increased to near-wild-type levels, due to (a) compensatory mutations in genes leading to constitutive high-level expression of the PhoE porin or (b) mutations in *hfq* and *chiX* genes that disrupted Hfq-dependent small RNA regulation, causing overexpression of the ChiP porin. These findings may explain why porin composition is often altered in resistant clinical isolates, thereby providing new insights into how bypass mechanisms may allow genetic adaptation to a common multidrug resistance mechanism.

# **Measurements of fitness costs.**

We (WP7) measured the fitness costs of mutational resistance mechanisms during growth under various types of *in vitro* and *in vivo* conditions and in different genetic contexts. Mutations were introduced into different bacterial species and clones to assess the impact of species and clone characteristics on fitness costs. Growth characteristics have been measured in single culture and during competitions between susceptible and resistant strains. Similarly, we have analysed potential epistatic effects between different types of resistance mutations by combining them in all possible combinations. Our main findings are the following:

- a) We demonstrated that the fitness and resistance effects of any given resistance mutation is largely independent of the genetic background used. Thus, when introducing the same resistance mutations (*rpsL*-streptomycin resistance, *rpoB*-rifampicin resistance, *fusA*-fusidic acid resistance and *gyrA*-fluoroquinolone resistance) into four different Salmonella strains that vary in their DNA sequence, no differences could be seen with regard to how much fitness was reduced and how much resistance was increased. This indicates that epistatic effects are not very strong for these chromosomal resistances.
- b) We analyzed potential epistatic effects between five different types of resistance mutations by combining these five mutations in all possible combinations. We observed few cases of strong epistatic interactions for these mutations and instead their combined fitness effects were largely additive.
- c) We examined the fitness effects of mutations that confer resistance to tigecycline, mecillinam and colistin. In general, these resistance mutations conferred a reduction in fitness ranging from a few percent up to 50%.
- d) We studied the fitness costs and distributions of different transposons carrying the *vanA* gene, contained in different plasmids (variety of genetic backgrounds) in a well-defined collection of vancomycin resistant *Enterococcus* responsible for hospital outbreaks around the world. Our results suggest that various Tn*1546* variants could spread in their plasmid vehicles across bacterial populations, but only a limited number of hosts ensure their stable maintenance and that differences in fitness might explain the particular association of particular genetic configurations with particular clones and species.

# **Evolutionary potential of low-level resistance genes.**

To assess the evolutionary potential of resistance genes we (WP3) identified sets of genes conferring lowlevel resistance to either fluoroquinolones or beta-lactams. The two drug classes differ significantly in the ability of chromosomal mutations to lead to high-level resistance phenotypes. In these studies we focused on two main parameters that would influence the risk of such low-level resistance genes to become highlevel resistance genes. First, we assessed the stability of plasmids expressing low-level resistance genes in absence of antibiotic selection pressure. Second, we assessed the extend to which evolution of the plasmids with low-level resistance genes could contribute to high-level resistance phenotypes in a host strain subjected to antibiotic selection pressure.

# **Stability of low-level resistance genes.**

#### *Stability of low-level flouroquinolone resistance genes*

Using metagenomic functional selections we identified 20 novel *qnr* genes conferring low-level resistance towards the flouroquinolone antibiotics. The stability of 20 plasmids containing genes encoding different Qnr-like proteins was analysed. Eight of the plasmids where shown to be unstable, indicating that fitness costs associated to the presence of plasmid-encoded resistance genes are allele specific. The other 12 genes were stable, which and do not produce relevant fitness costs.

#### *Stability of low-level beta-lactam resistance genes*

In contrast to the *qnr* genes, previous work have shown that plasmid-encoded beta-lactamases can evolve *in vivo* under antibiotic selective pressure. To assess this possibility we selected genes from functional metagenomic libraries conferring low-level resistance to beta-lactams. We focused on clones containing putative beta-lactamases, PBPs or hypothetical proteins. We did not take into consideration clones containing genes encoding regulators as MarA, since the role of these genes on resistance depends on the presence in the host genome of the genes they regulate. In total we focused on 6 low-level beta-lactam resistance genes. Adaptive evolution of the lineages containing each of the 6 resistance genes in absence of selective pressure showed that the plasmids were very stable. This data indicates that the fitness cost of expressing the low-level beta-lactam resistance genes was low.

#### **Evolvability of low-level resistance genes**

#### *Potential of low-level flouroquinolone resistance genes to become high-level resistance genes.*

The 12 stable *qnr* genes from described above were subjected to evolution by sequential sub-culturing in increased concentrations of quinolones. After 25 days (around 200 generations), the MICs to quinolones of the strains carrying these plasmids increased by several fold. The evolved plasmids were extracted and used to transform a wild-type strain. The aim was to establish whether the mutations leading to quinolone resistance were in the low-level resistance genes or genomic. In parallel, the genes coding the Qnr-like proteins present in the evolved plasmids were sequenced. None of the evolved plasmids increased the MICs for quinolones to the level observed in the evolved strains when they were re-introduced in a wildtype strain suggesting that the mutations leading to quinolone resistance in the evolved strains were chromosomally encoded. In agreement with this hypothesis, none of the sequenced *qnr* genes presented any relevant mutation. Our results indicate that the risk that Qnr-like elements evolve towards high-level quinolone resistance is not high, likely because mutations at topoisomerase genes are easily selectable in the presence of quinolones. This result is in agreement with information concerning QnrA and SmQnr published at (Sanchez, M. B.; Martinez, J. L., PLoS ONE 2012, 7, e35149). If this hypothesis holds true, the different alleles of *qnr* genes currently present in human pathogens should not have evolved under quinolone selective pressure at clinics, but rather represent different acquisition events. Following this work we analysed the structure of plasmids containing *qnrA* and *qnrB* genes so far present at public databases. Our results indicate the different *qnrA* or *qnrB* alleles have polyphyletic origins suggesting that the *qnr* genes currently present in the plasmids of human pathogens are not the result of evolution under antibiotic selective pressure in clinics.

#### *Potential of low-level beta-lactam resistance genes to become high-level resistance genes.*

The 6 low-level beta-lactam resistance genes where introduced into an *E. coli* strain defective in *mutS* and *ampC* and evolved in ampicillin and cefotaxime. In all cases, strains evolved to acquire high-level resistance to the drug used. For the ampicillin-evolved clones, plasmids were rescued and an *ampC* minus, but *mutS* proficient strain was transformed with both the original and the evolved plasmids. In all cases, the evolved plasmid conferred higher resistance than the original one, indicating that mutation was in the plasmid. Noteworthy, the mutations present in the plasmid were synonymous and did not alter the amino acid sequence. This could indicate that evolution of elevated beta-lactam resistance could result from different protein expression levels resulting from differential codon usage.

As indicated above evolution and emergence of antibiotic resistance is not only the result of mutations but also of acquisition of resistance genes importantly through transfer of plasmids carrying resistance genes. In WP6 the dynamics of antibiotic resistance transfer within bacterial populations was studied and novel methods were developed to quantify resistance transfer under controlled conditions in bacterial communities. To achieve these objectives, and according to EvoTAR technical annex, work was divided in six specific tasks.

## **Characterization of environmental factors that affect plasmid stability and propagation.**

We compared the conjugation kinetics of representatives of five prevalent plasmid groups (after constructing derivatives labeled with fluorescent protein reporters), in an effort to gain insight about potential differences that can explain the prevalence of different plasmid groups. The repressed (wild type) versions of IncF and IncI plasmids, although very prevalent in enterobacterial populations (specifically *E. coli*), showed reduced conjugal infectivity. This result suggests that alternative fitness components are important in determining the success of these plasmids. One potential compensation probably originates from the decreased burden that repressed plasmids cause on the host bacterial population. More work needs to be done to define these compensatory effects more rigorously. In general, it can be concluded that each of the five tested plasmid groups has some specific properties that affect its conjugation kinetics and, thus, its infectivity of susceptible recipient populations. These differences can contribute to a rational explanation for the prevalence of different plasmids in enterobacterial populations. Obviously, infectivity rates are not the sole cause to explain differential prevalence. Nevertheless, our results show that infectivity rates can represent a significant contribution and, perhaps more importantly, that each plasmid group shows differential parameters that are experimentally testable. When combined with the analysis of differential stability in different hosts under different conditions, this type of analysis can lead to a sort of "specification sheet" that defines a set of relevant parameters for each plasmid group.

# **Identification of signals that trigger conjugation**

Studies on plasmid conjugation kinetics led to a general hypothesis of the role of transcriptional overshooting in plasmid conjugation and genome rebooting. The main achievement was the elucidation of plasmid R388 transcription control system, and the finding that transcriptional overshooting probably provides the main systems-level control of plasmid conjugation. The signals that trigger conjugation are thus endogenous, rather than environmental (fitness changes of recipients themselves after conjugation cause the observed variations in apparent conjugation frequencies). Another significant result was that plasmid R388 is able to infect a recipient population based on privileged donor multiplication. The reason for this effect is not known. It seems plasmid R388 can cause some detrimental effect on the recipient population. This is comparable to the effect of bacteriocin, although R388 is not known to code for any bacteriocin. Other, more complex, causing mechanisms can be envisaged (conjugation-induced killing, etc.). Besides, the cost of gene amplification on fitness and stability was quantified for an IncI plasmid (Enterobacteriaceae, γ-proteobacteria) and for plasmids of different families (Enterococci, Firmicutes).

# **Quantitation of fitness costs incurred by plasmid carriage.**

The results obtained, both in Gram-positive as well as Gram-negative bacteria, show that fitness costs of carrying resistant plasmids are strongly dependent on the particular plasmid, the genetic host and the growth conditions. Thus, fitness costs vary between low (undetectable) to up to 20% or even more, depending on conditions. Comparative genomics was used to analyze compensatory mutations in evolved strains. Results showed that adaptive mutations and indels occur either in the chromosome or in the plasmid. Several chromosomal mutations and indels were associated to genes involved in key cellular functions. The majority of the changes observed in evolved plasmids were due to deletions associated with key functions (replication, conjugation and maintenance) or accessory genes (antibiotic resistance, plasticity – IS elements). These genome alterations might be responsible for the fitness differences observed between evolved and non-evolved strains. Preliminary data show that evolved strains show improved fitness compared to their non-evolved counterparts. In a particular case studied in detail, when antibiotic selective pressure was applied, rapid amplification of the resistance-conferring gene to high copy numbers occurred. When selection was relieved, the amplified array rapidly disappeared. This mechanism provides a rapid and reversible adaptive mechanism for bacteria to increase their resistance under strong selective pressures. Because of the high instability of the amplified arrays, it is likely that the clinical microbiology laboratory will typically miss the importance of this type of mechanism. The complexity of factors involved in plasmid fitness cost is also supported by results obtained from studies

performed with two epidemic plasmids with a major role in the emergence and dissemination of CTX-Mtype ESBLs in Bolivia. Despite the natural history of those plasmids (the IncA/C plasmid was displaced by the IncI1 plasmid), in vitro experiments performed in diverse experimental conditions did not provide evidence explaining such epidemiological change. Ongoing studies might contribute in identifying predictor markers for plasmid success.

# **Evolution of plasmid host-range**

Important observations made by other authors (e.g., the group of Eva Top, University of Idaho) indicate that important changes in plasmid host range occur easily within a given backbone and depend on quick adaptive mutations. The host-range itself is not specific of a given plasmid group, but changes widely among very similar plasmids. These results, which were published between the writing of EvoTAR and the beginning of work, affected the type of experiments that were carried out by EvoTAR members and, in general, diminished the interest in finding specific mutations that affect host range. Nevertheless, some effort was put to analyze the mechanisms of evolution of plasmid host-range in specific cases. When using Gram-positive plasmids, results showed that multireplicons can be a strategy to broaden the host range of antibiotic resistant plasmids among opportunistic pathogens sharing common habitats. When using ColE1-like plasmids and studying their adaptation to *E. coli*, it was found that plasmids rapidly evolve increased stability, by either point mutations or IS-element insertions. When several plasmids coexist, the overall burden is the additive value of the individual burdens. As anticipated, the problem with these studies is the difficulty to generalize, due to the plasmid individuality in the changes that affect host-range.

# **The natural history of plasmid adaptation**

As a general approach to the problem, we developed a plasmid reconstruction method for Illumina whole genome datasets, called PLACNET. PLACNET allowed us to achieve a giant leap in the analysis of natural history of plasmid adaptation, by providing massive data for comparative genomics. The implications of this method in the analysis of the natural history of plasmid adaptation (both in Gramnegatives and Gram-positives) is exemplified in several publications of high impact. Among other results, we select two important conclusions as examples of the applications of this bioinformatics technology: the turbulent plasmid flux in *E. coli* ST131 (a measure of the speed of plasmid evolution compared to core-genome evolution), the transfer of epidemic plasmids from animal to human *E. coli* isolates as the cause for dissemination of ESBLs, and the characterization of the Firmicutes plasmidome. Besides the grand scenario, EvoTAR also tackled the natural history of specific examples of antibiotic resistance dissemination. Analysis of *E. coli* plasmids encoding cephalosporin resistance from various origins, indicates that spread of ESBL-encoding plasmids occurs by HGT among *E. coli* lineages. The finding of nearly identical plasmids in different Enterobacteriaceae and in epidemiologically unrelated individuals suggests that such plasmid lineages possess traits involved in host adaptation (e.g., IncHI1 plasmids carrying a sugar metabolic element likely enhancing *E. coli* fitness in the equine gut). Findings show that specific plasmid lineages contribute to global ESBL spread within host species with limited overspill between hosts. With respect to specific studies on Gram-positive plasmids, the persistence of VanA-type VanR 15 years after the ban of avoparcin was shown. *vanA* was mainly linked to a specific plasmid lineage that was non-typeable and did not carry genes conferring resistance to antimicrobials used in poultry production. Findings suggest adaptation of such plasmid to *E. faecium* in the avian gut. An additional study comparing clinical and avian VRE isolated in Denmark is ongoing. Preliminary results suggest that vanA-encoding plasmids in human and avian *E. faecium* are not directly linked.

# **Analysis of antibiotic resistance gene transfer in controlled environments**

We attempted to construct a device that could be used as a universal conjugation sensor by taking advantage of the SOS-response provoked by conjugation. All attempts failed, so we are still trying today to achieve this goal, which we consider will be a useful addition for the antibiotic-resistance analysis toolbox. Therefore, we returned to more classical means for studying antibiotic-resistance transfer in controlled environments mimicking natural ecosystems. We developed two such systems: a freshwater microcosms and a mouse societal model. In both systems, proof of principle was obtained that we can track antibiotic-resistance transfer. This will allow detailed studies of the effects of system and environmental parameters on the dissemination of antibiotic-resistance. Experiments to test the efficacy of conjugation inhibitors on both systems are underway.

A major objective of the EvoTAR project in general, and of WP7 in particular, was to develop generic and predictive models that allow a detailed description of the within-host dynamics and between-host dynamics of antibiotic resistance modules (genes, genetic elements, clones) and that will quantify the probability and rate of emergence and spread of resistance-conferring genes/mutations under various environmental conditions, different selective pressures and in different genetic backgrounds. This modelling-based approach is essential for the prediction of the risk that a given antibiotic resistance gene may successfully spread to pathogens and thereby contribute to future resistance problems and how changes in the selective pressures influence rates of spread of antibiotic resistance at the population level. To provide parameter values for the various models, experiments were performed to determine fitness of various single and multi-drug resistances in combination and in different genetic backgrounds. Below we summarize the major findings from this work under four different headings—within host level models, between host level models, mixed level models and measurements of fitness costs.

#### **Within host (individual) level models**

*In silico* pharmacokinetic-pharmacodynamic (PKPD) models can be developed based on data from *in vitro* time-kill experiments and can provide valuable information to guide dosing of antibiotics. We developed a mechanism-based *in silico* model that can describe *in vitro* time-kill experiments of *E. coli* wild type, and six isogenic mutants, exposed to ciprofloxacin and to identify relationships usable to simplify future characterizations in a similar setting. The developed model includes susceptible growing bacteria, less susceptible (pre-existing resistant) growing bacteria, non-susceptible non-growing bacteria and non-colony-forming non-growing bacteria. A common model structure with different potency for bacterial killing for each strain successfully characterized the time-kill curves for both wild type and the six  $E.$  *coli* mutants. Our results show that the model-derived mutant-specific  $EC_{50}$  estimates were highly correlated with the experimentally determined MICs, implying that the *in vitro* time-kill profile of a mutant strain is predictable by the MIC alone based on the model.

# **Between hosts (population) level models.**

*Efficient national surveillance for healthcare associated infections.* A general model framework for the spread of a disease over a network of connected nodes was developed. To validate this model and illustrate the potential use of it, we simulated the transmission of a novel HCAI that spreads predominantly by direct patient movement between Scottish hospitals as a result of patient movements, i.e. with little or no transmission in the community, and extend it by comparing existing surveillance programs with a (putative) optimal program to see if, with easily acquired information on the network of patient transfers, existing national surveillance schemes can be made more efficient. The model enables us to prioritise hospitals for inclusion in a laboratory surveillance system and thereby to address two key questions: 1) What is the optimal distribution of surveillance effort across hospitals and is the current system maximally efficient; and 2) Would there be benefits from increasing or decreasing the number of hospitals engaged in surveillance? Our analyses show that the current surveillance system, as it is used in Scotland, is not optimal in detecting novel pathogens when compared to a gold standard. However, efficiency gains are possible by better choice of sentinel hospitals, or by increasing the number of hospitals involved in surveillance. Similar studies could be used elsewhere to inform the design and implementation of efficient national, hospital-based surveillance systems that achieve rapid detection of novel HCAIs for minimal effort.

*Multistrain network model.* The first model was extended by incorporating the possibility to simulate the spread of multiple "strains" on a network, leading to a generic model to study the spread of different resistances (genes or plasmids) between hosts. To illustrate the potential use of this model a similar approach was used as with the first model. With this model it is possible to predict where potential hotspots are for the emergence of multidrug-resistant strains. To illustrate this, we calculated the frequency of each individual hospital to be the first hospital to be co-colonised with the two strains over 20000 simulations. The results show that some hospitals are (on average) more frequently the first to be

co-colonised then others. Identification of these hospitals show that mainly (large) teaching hospitals are hotspots for the emergence of multidrug-resistant strains (there are seven teaching hospitals in Scotland and the seven most frequently co-colonised hospitals are those seven hospitals, followed by large community hospitals).

## *The effects of population structure on the long-term prevalence of antibiotic resistance.*

To study the impact of population structure on the long-term prevalence of an antibiotic resistant strain a novel model was developed. The results of this model show that having a batch-structure might be beneficial for keeping antibiotic resistance levels low. The main finding of this model is that having a continuous structured population (like a human population) might lead to a higher long term prevalence of an antibiotic resistant strain than a batch-type structured population if all other parameters are kept the same and that a batch-structure population is much harder to invade than a continuous birth-death population, i.e. to reach comparable prevalence levels between the two population structures the "spontaneous creation rate" of an antibiotic resistant strain needs to be much higher in the batch-structure population compared to the continuous birth-death population.

*A hospital-level risk factor analysis of Staphylococcus aureus bacteraemia in Scotland.* The outcomes of the first model (specifically the connectivity of hospitals) were also used in a hospital-level risk factor analysis for *Staphylococcus aureus* bacteraemia cases in Scotland. The aim of this study was to identify risk factors for the presence and rate of MRSA bacteraemia cases in Scottish mainland hospitals. Specific hypotheses regarding hospital size, type and connectivity were examined. In Scotland, although hospital size is a significant predictor of the presence and rate of MRSA, it does not fully explain all the observed variation among hospitals. In this study we found that in Scotland, there is a certain level of connectivity above which the majority of hospitals, regardless of size, are positive for MRSA. Higher levels of MRSA are associated with the large, highly connected teaching hospitals with high ratios of patients to domestic staff.

*Antimicrobial prescribing and its relationship with antimicrobial resistance in MRSA.* The specific aims of this study were two-fold. Firstly, to examine spatial and temporal trends in Scottish primary and secondary care prescribing rates. Secondly, to investigate whether or not there were any associations between primary or secondary care prescribing rates and antibiotic resistance in the MRSA population. To address this, there have been calls for improved antimicrobial stewardship to better regulate drug usage, as well as improved surveillance, monitoring and regulation. Firstly, it was found that antibiotic usage of several antimicrobials increased and the rate of this increase should be monitored to prevent extreme over-use and drugs potentially becoming obsolete. Secondly, the rate of prescribing of different antimicrobials differed between HBs and over years which could be due to several factors but likely mirrors differing HB-specific prescribing guidelines but also represents a lack of consistency in treatment. Thirdly, resistance was found to be associated with prescribing rates for three antimicrobials over this study period, although there are also likely to be other factors contributing to resistance (for example historic prescribing).

*Global disease burden due to antibiotic resistance – state of the evidence.* The absence of comprehensive and reliable estimates of the global health burden due to antibiotic resistance makes it difficult to assess trends and harder to justify the allocation of adequate resources to deal with the problem. Quantification of the burden of resistance requires data on the incidence of clinical conditions appropriately treated with antibiotics, the frequency of treatment failures due to resistance and their impact on clinical outcome. Treatment failures in turn depend on the level of resistance in the aetiological agent to the antibiotic used. These data are not easily obtained, as illustrated by a case study of neonatal sepsis. One obstacle is that global health statistics as currently collected do not provide the necessary information. Improving this situation will require changes to the ways in which global health statistics are collected. The primary benefit will be more accurate assessment of the global disease burden due to antibiotic resistance and its forward trajectory, helping make the case for investment in combating the problem.

*The utility of Whole Genome Sequencing of* Escherichia coli *O157 for outbreak detection and epidemiological surveillance.* This study assessed the utility of whole genome sequencing (WGS) for

outbreak detection and epidemiological surveillance of *Escherichia coli* O157, and the data was used to identify discernible associations between genotype and clinical outcome. The results show WGS data can provide higher resolution of the relationships between *E. coli* O157 isolates compared with MLVA. The method has the potential to streamline the laboratory workflow and provide detailed information for the clinical management of patients and public health interventions.

*Intercontinental exclusion of community-associated MRSA subtypes with distinct antibiotic resistance gene profiles.* We investigated the evolution and global spread of ST59, a pandemic, communityassociated clone of *S. aureus*, which is found globally and is a major cause of skin and soft tissue infections in south-east Asia. We showed that two distinct ST59 clades emerged independently, one in Taiwan and the other in the USA. Whilst both clades also contained sequences from Australia and Europe, no exchange of strains between Taiwan and the USA was observed in either direction. We also found that strains in the Taiwan clade possessed a greater number of both antibiotic resistance and virulence determinants than the USA clade. Using growth experiments in the laboratory, we demonstrated that ST59 strains from the Taiwan and USA clades were able to out-compete USA-300, the dominant community-associated strain in the USA. Our findings are consistent with the hypothesis that differences in antibiotic usage in the USA and Taiwan, and competition with other *S. aureus* strains including the dominant community-associated strain in the USA (USA-300), could explain the lack of transfer of ST59 between Taiwan and the USA.

*Novel transmission model.* We have established an innovative, cheap, and reproducible experimental model of transmission of information about antibiotic resistance using two large (>50 individuals) populations of the cockroach *Blatella germanica* placed in two compartments, one with frequent antibiotic exposure (mimicking a hospital), and the other with minimal antibiotic exposure (the community), with a certain rate of migration between the environments. This experimental model can be used to address how population structure, density, migration, antibiotic pressure etc. influences the transmission of resistant bacteria in a population and between different environments and compartments.

# **Mixed level models.**

We developed a computational multi-scale modeling to connect performances arising at distinct scales. These models are known as nested or embedded models and have been used to address specific questions involving within-host dynamics enclosed in a model of between-host epidemiological scenarios. This approach requires a nested model because the different units involved in resistance are nested units of selection across distinct scales of the subcellular, cellular and supra-cellular environmental levels of the ecosystem, including communities of hosts. Thus, any alteration of the carriers of any particular resistance trait, or its mechanisms of variation and mobilization (mutation, recombination, transposition, lateral gene transfer, migration) may influence the dynamics of other units of higher and lower hierarchy and thus have evolutionary and/or ecological consequences on a bacterial population. The difficulty for modeling this type of ecosystem scenarios with nesting has been an important limitation to convincingly study processes of AR evolution, but exciting new opportunities have recently arisen from a natural computing formalism inspired on the structure and functioning of biological cells, called membrane computing. Membrane-computing considers that any biological system is a hierarchical construct where the flow of materials can be interpreted as computing processes. In particular, membrane computing offers a versatile framework known as P-system that consists of a hierarchical membrane structure of nested compartments (regions) where multisets of objects are located and can move across the resulting "membranes" and evolving according to a finite number of given rules. Using this approach, we developed a new computational approach designed for computing at more than three levels of organization, subcellular, as genes or plasmids, cellular, and, supra-cellular through the software implementation of a simulator Antibiotic Resistance Evolution Simulator (ARES). ARES will facilitate predictive computational models on the potential trans-hierarchical response of antibiotic resistance to particular interventions in specific scenarios.

Two different approaches were pursued in EvoTAR WP8 to interfere with the selection and dissemination of antibiotic resistance. The first approach aimed to reduce the free antibiotic concentration in the gut, thereby reducing the risk antibiotic resistance selection. The second approach aimed at interfering with conjugative transfer of antibiotic resistance genes.

## **Reducing antibiotic concentrations in the gut.**

The vast majority of orally administered antibiotics are only partially absorbed into the blood reaching the intestinal tract, and for some of them, a significant part of the administered drug remains intact before reaching the colon. A similar phenomenon occurs for parenterally administered antibiotics that are recycled, via the hepatobiliary route, from the blood into the small intestine. Thus, for both oral and parenteral antibiotics, active residues reach the colon at doses that are lethal for most commensal bacteria. These residues thereby provoke serious collateral damage amongst the intestinal microbiota of patients: their gut microbiota balance is disturbed; several bacterial populations are erased whereas other strains proliferate. Antibiotic-treated patients' microbiota will need several months to recover. This phenomenon brings along harmful consequences as antibiotic resistant bacteria are selected in the gut. It indeed plays a key role in the onset of *Clostridium difficile* infections (CDI) by allowing resistant *Clostridium difficile* bacteria to outnumber other intestinal bacteria, causing very painful and deathly diarrhea. In addition, selection of resistant bacteria in the gut also leads to increase global antibiotic resistance as resistance genes are passed on to the many strains of bacteria present in the commensal flora via exchange mechanisms.

In EvoTAR, partners in WP8 contributed to the development of two microbiota-focused interventional products: DAV132 for human health and DAV133 for animal health, helping innovation to translate into marketed products.

DAV132 is a groundbreaking approach whereby a highly efficient adsorbent, with a specific intestinal delivery technology, can be co-administered with virtually any antibiotic. DA132 is developed to prevent CDI, to avoid the havoc that antibiotics wreak upon the gastrointestinal tract as they are eliminated from the body, and eventually dramatically reduce antibiotic resistance development. DAV132 inactivates the unwanted fractions of the antibiotic that reach the lower digestive tract and remediates it through stool, after the antibiotic has been systemically absorbed to fight infections. In EvoTAR, clinical batches of DAV132 were manufactured in compliance with all standard regulations and then used in a clinical study performed in France with 44 healthy volunteers. The study was randomized with 4 groups: 14 volunteers received oral moxifloxacin (a widely used quinolone antibiotic), 14 volunteers received moxifloxacin associated with DAV132, 8 received DAV132 alone and 8 received no treatment at all. The study demonstrated that, in humans, DAV132 was well tolerated. The metagenomic analysis performed in EvoTAR also illustrated that DAV132 protects humans from fecal microbiome disruption after oral moxifloxacin treatment, without decreasing the efficacy of the antibiotic (no change in the plasma concentration). It is the first time a product successfully achieves this result and it opens the way for a novel use of antibiotics with much less harmful consequences. The work of the EvoTAR program was immensely helpful in the clinical development of DAV132. Upon these foundations, the development work will continue with the hope DAV132 product will be introduced on the market in some years.

DAV133 relies on a similar mechanism of action to preserve animals' microbiota from antibiotic disruption and particularly avoid the selection of resistant bacteria. The medical interest of this approach is to limit the carriage of resistant bacteria in livestock and companion animals to thwart the transfer of antibiotic resistant bacteria from animals to humans. In this regard, DAV133 is much in line with the One Health approach stating the importance of taking into account animal health questions to improve human health. In the context of the EvoTAR program, DAV133 was optimized for animal use (special formulation, specific administration scheme), manufactured and then tested in an animal clinical study. This also vastly contributed to the development of the product.

## **Interfering with conjugative transfer of antibiotic resistance genes.**

Transfer of antibiotic resistance plasmids from one bacterium to another relies on the use of different conjugation systems that were identified as potential targets for the development of inhibitory treatments that could reduce the spread of antibiotic resistance. Instituto Biomar participated in the Evotar project with the aim of establishing the proof of concept for this approach, focusing on the inhibitor of conjugation AD0149, previously identified in collaboration with other Evotar partner, Universidad de

Cantabria. Loss of the conjugation inhibitory properties of the compound upon purification over 95% indicated that the inhibitory activity was associated with minor components of the original sample that could not be identified despite strong efforts in that direction. This barrier resulted in an intense screening campaign, that led to the identification of several inhibitors, being the family of the Tanzawaic acids the most potent. These results will be published in the next months.

Characterization of the new inhibitors highlighted the need for chemical modification of the structure to improve water solubility, a very relevant property, as the intended system to evaluate the compound was an aquatic microcosm, and also for the potential application of the inhibitors in aquaculture. Finally, through characterization of the modified compounds and some additional candidates in the screening system, a demonstration of the capacity of some inhibitors of conjugation of blocking transfer of antibiotic resistance in a bacterial population was provided. These results, obtained in collaboration with Universidad de Cantabria, support the potential application of conjugation inhibitors in the control of antibiotic resistance spread. Some challenges, related to the potency requirement, cost and regulatory issues will need to be resolved to allow the establishment of conjugation inhibitors as real tools in the control of the resistance problem. Future work with models resembling use of antibiotics in hospitals or farming settings will help resolving these pending challenges.

#### **Summarizing conclusions of the main results**

The aim of EvoTAR was to characterize humans, animals and environmental reservoirs of antibiotic resistance genes to study the dynamics within and interactions between these reservoirs. In five figures, we have tried to summarize important findings of the EvoTAR with respect to its aim. In these figures numbers indicate data generated within the EvoTAR project. Below every figure this is explained in more detail. Figure 10 summarizes all findings that are detailed in figures 6-9.



Figure 6. Studied reservoirs and observed links between reservoirs of antibiotic resistance genes, mobile genetic element and strains. Red coloured and numbered squares refer to data generated within the EvoTAR project.

1. Only few cases of *E.coli* or *E. faecium* clones shared by humans (farmers) and food animals, pigs can be documented (WP2). However, pigs and humans share a number of antibiotic-resistance

genes as detected by gene capture, and pigs are highly enriched in a high diversity of resistance genes (WP4). Clones and resistance genes are frequently maintained within the pigs. Metatranscriptomic datasets from intestinal microbiota from pigs demonstrate that genes are locally expressed.

- 2. Whole genome sequence studies reveal low-likelihood of recent transmission of ESBL-*E.coli* clones between poultry and humans (WP2). Also *vanA***-**encoding plasmids in human and avian *E. faecium* are not directly linked (WP6). However, some plasmids (IncK plasmid backbones) are shared by avian and human Enterobacteriaceae (WP6); in general promiscuous plasmid transfer among bacteria of different reservoirs is possible (WP2) but infrequent.
- 3. Isolates from marine environments are infrequently resistant, but some of them are multiresistant. In soil there are bacteria with resistance againt vancomycin and also with resistance against carbapenems that can be expressed in *E.coli*; there is no evidence of transfer to human isolates (WP5).
- 4. Isolates from water bacterioplankton, sponges, water sediments might contain resistance genes that are locally expressed, suggesting a role in ecological adaptation unrelated with antibiotic resistance (WP5). Transfer to humans was not documented.
- 5. Bacteria of the predominant commensal taxons in the intestinal human microbiota contain a wealth of genes (intrinsic resistome) able to provide resistance to the bacterial host (WP4). These genes are locally expressed under antibiotic exposure, which explain the maintenance of richness during therapy (WP1), but are very unfrequently transmitted to significant bacteria for public health. Low-level antibiotic resistance genes are frequently detectable (WP3), but not clear evidence of evolution to high-level resistance was found.
- 6. Significant bacteria (significantly pathogenic) for human health, as *Enterobacteriaceae* or *Enterococcus* are present in <1% of the microbiome population. These organisms have genes of the intrinsic resistome (WP2), and are the populations with a higher density of significant resistance genes (WP4). Under therapy these populations reach high densities significantly increasing the number of resistance genes in the human intestinal microbiome (WP1).
- 7. There is no clear evidence of significant rate of transfer of these genes from potentially pathogenic populations to majority taxons, but that might happen (genetic exchange communities, for instance *Enterococcus-Lactobacillus-Clostridium*) (WP4).
- 8. The increase of the number of resistant organisms resulting from antibiotic selection facilitates the spread of plasmids containing antibiotic-R, among closely-related clones or taxons (WP5 and 6).
- 9. There is a solid link of plasmids and host lineages, suggesting epistatic genome-plasmid robustness and/or contribution to local adaptation (WP6, WP7).
- 10. Some inhibitors of plasmid transfer by bacterial conjugation, as tanzawaic acid derivatives, have been found (W8) which might limit plasmid spread under antibiotic selection.
- 11. Short-term antibiotic exposure do not produce a significant enrichment of antibiotic resistance genes in the microbiome (WP1). Very long-term exposure significantly increases overall level of resistance genes, even if the richness is decreased. Standard (hospital) antibiotic exposure do not decrease richness, but the number of resistance genes is increased, probably because of the selection of minorities (WP1, WP4). Reduction of antibiotic exposure can be achieved by adding novel charcoal-based products aiming to adsorb residual antibiotics in the colon while not interfering with the duodenal/jejunal absorption (WP8).



Figure 7. The human resistome in the absence of antibiotic exposure.

- 12. Upon discontinuation of antibiotic exposure, the intrinsic resistome of the majoritarian taxons is less active; the minority populations returns to their normal abundances, with decrease of the total density of detectable antibiotic resistance genes (WP3).
- 13. Some mobile genetic elements encoding resistance traits imposes a very small fitness cost to their hosts, and the consortium is stably maintained over time (WP7) and suggest influences on host ecological adaptation (WP6).
- 14. Starting after one month, after three-six months of antibiotic discontinuation the populations of the intestinal microbiome reach the equilibrium point, returning to the original population structure (WP1).



Figure 8. Maintenance of antibiotic resistance genes and plasmids.

- 15. Most clones and species harboring resistance genes and plasmids are stably maintained even without antibiotic exposure, suggesting a resistance-host long-term adaptation (WP4, WP3, WP7).
- 16. Fitness costs of harboring (and expressing) particular mobile genetic MGE elements might remove some resistant organisms from the microbiome (WP7).
- 17. Fitness costs of mutations in resistance genes are variable and depends on the host genetic background (WP7).
- 18. Some MGE are able to spread but only within a limited number of hosts (WP7, WP4).



Figure 9. Dissemination of antibiotic resistance

- 19. Antibiotic selection and the resulting epidemic-endemic spread of resistant clones increases the variability and adaptation of hosts and is a major factor in the enrichment of resistance genes (WP2). Small antibiotic concentrations might be selective and amplify resistance plasmids and genes (WP2).
- 20. A major factor contributing to the expansion of resistance genes is the transfer of colonized patients between *connected* hospitals (WP7).



Figure 10. Summary of reservoirs and observed links between reservoirs of antibiotic resistance genes, mobile genetic element and strains in the presence and absence of antibiotic induced selective pressure, and describing aaintenance of antibiotic resistance genes and plasmids and dissemination of antibiotic resistance. The used three colour codes in the figure indicate strong/weak/negligible flow or source of resistance genes.

# **Potential impact, main dissemination activities and exploitation of use**

Novel approaches to identify antibiotic resistance determinants (ARDs) from complex, metagenomic datasets in WP1 allowed the construction of a large list of the ARD from the human gut genome. This lays grounds for detailed investigations of the resistome dynamics under different conditions – active treatments with antibiotics or passive exposure due to residues in the food or environment. These investigations can open avenues to reducing the overall load of ARDs, by adjusting treatments and implementing practices to reduce the residues in the environment. Beyond humans, reservoir of the ARDs in domestic and farm animals could be followed and the transfer from the zoonotic sources to humans explored.

The methodology developed to identify ARDs, based on structure modelling, has a potential to be also applied to identification of other functions of interest in the human gut microbiome that could be associated to health and disease. Indeed, a majority of the millions of genes that are known have no clearly identified function. Our comprehension of the overall microbiome function is greatly hampered by this lack.

Studies in WP2 have uncovered multiple pathways that can lead to the evolution of antibiotic resistance and successful clones of nosocomial pathogens. These findings have had considerable impact and may

contribute to novel screening methods and/or interventions aimed to detect and stop the spread of highrisk antibiotic-resistant bacteria. Some of the most noticeable pathways to impact for EvoTAR WP2 are outlined below.

- Multi-drug resistant clinical *E. faecium* isolates form a specific sub-population which has emerged from human commensal and animal strains. Several genetic elements are present in clinical *E. faecium* strains while being absent in human commensal and animal strains. These genes may be used to develop rapid PCR-based strategies to distinguish clinically relevant *E. faecium*, which may have the potential to spread rapidly in hospitals, from strains that are not adapted to cause outbreaks among hospitalized patients.
- The study on the dissemination of cephalosporin-resistant *E. coli* strains from poultry, chicken meat and hospitalized patients provided an important correction on previous studies, which used low-resolution traditional typing methodologies to study the relatedness of strains from these different reservoirs. Our study could not substantiate previous claims on direct transfer of resistant *E. coli* strains from chickens via meat to humans but highlighted the important role of identical plasmids spreading between *E. coli* strains from different reservoirs. As plasmids can spread promiscuously between bacteria from different species, this may considerably complicate containment of resistance plasmids from diverse reservoirs. In addition, this study highlighted the superiority of whole-genome sequencing for bacterial typing, compared to other methods like MLST and PFGE.
- While many research groups appreciate the usefulness of whole-genome sequencing for bacterial typing, there is often a hurdle on the implementation of this technique due to the requirement of bioinformatics analyses. To open up whole-genome sequencing as a typing approach for laboratories without dedicated bioinformatics support, a core genome MLST scheme for *E. faecium* was developed, which can be used with minimal bioinformatics expertise.
- Several studies in EvoTAR have highlighted the rapid emergence of colistin resistance in *K. pneumoniae* through multiple evolutionary trajectories*.* Worryingly, resistance to colistin can come at no detectable fitness cost. Our findings should lead to guidelines that minimize the nonessential use (e.g. in farming and, potentially, in prophylactic antibiotic therapies) of colistin, an antibiotic of last resort, to minimize the emergence of resistance.
- Studies in WP2 have expanded on previous observations that exposure to low levels of antibiotics can lead to high-level resistance. This mechanism may lead to the selection for resistance among bacterial populations in the environment. These findings may lead to interventions aimed at minimizing the release of antibiotics into the environment (e.g. through wastewater).

Dissemination of research data in WP2 was mainly performed through publications in scientific journals and presentation at microbiology conferences, including large international meetings like those organized by the Federation of European Microbiology Societies and the American Society for Microbiology. The majority of articles published in WP2 are available under a 'gold' or 'green' open access license, which has contributed to the visibility of the project.

The results of WP3 are expected to have a significant impact on the research field of antibiotic resistance. The impact here is both in the form of creating new and enabling technologies that accelerate research also beyond antibiotic resistance as well as building scientific knowledge about antibiotic resistance.

# **Enabling technologies:**

*Higher throughput functional metagenomics:* In WP3 we have developed several improved methodologies for performing high throughput functional selections. These methodologies (ParFUMS as well as its extension based on PacBio sequencing) have increased with throughput of functional metagenomics by 10-100 fold. Accordingly, we have in the EvoTAR project generated the largest dataset of antibiotic resistance genes which is equivalent in size to the current largest databases of antibiotic resistance genes (CARD maintained by McMaster University). These tools allow much more comprehensive characterization of resistomes also outside the clinical setting. Furthermore, these developments can be applied to the area of industrial biotechnology where functional metagenomics can be used to improve bioprocesses (Forsberg et at. 2015 AEM).

*New hosts for functional metagenomics:* We developed a methodology for construction of functional metagenomic libraries in *L. lactis* a Gram-positive organism. While we where not able to identify specific Gram-positive resistance genes using this approach, it still enables new paths for research within functional metagenomics, including the critical testing of expression bias for functional metagenomics.

*Rapid diagnostics for resistome dynamics:* We develop a methodology based on the Fluidigm Biomark system to interrogate a large set of clinically relevant resistance genes from any sample. With more testing and validation this approach could potentially be used as a rapid diagnostic for resistance genes in infecting pathogens.

# **Scientific knowledge:**

*Substantially expanded catalogue of resistance genes:* Through the EvoTAR project we have identified several thousand new variants of resistance genes using the improved functional metagenomic methods that we developed. This has enabled us to construct a better catalogue of resistance genes that will be made publicly available. Accordingly, we have uncovered more of the biological dark matter, e.g. the unknown resistance genes of the human gut microbiome.

*Improved understanding of resistome dynamics:* We have characterized the temporal dynamics of the resistome during ICU stay (including massive drug treatment). We have found that the resistome expands during hospitalization, both in terms of abundance of resistance genes as well as in terms of diversity of resistance genes. Furthermore, our findings suggest that the resistome during hospitalization is more enriched in resistance shared by human pathogens.

*Limited potential for evolution of low-level resistance genes:* We have tested resistance genes conferring low-level resistance towards both fluoroquinolone and beta-lactam antibiotics. Our results showed that there was a substantial interaction between the host genome and the low-level resistance gene that determined the evolutionary potential of the low-level resistance gene. For fluoroquinolones high-level resistance can more easily evolve through chromosomal mutations in *E. coli* compared to beta-lactams. Accordingly, fluoroquinolone low-level resistance genes cannot readily evolve into high-level resistance genes in such hosts. In contrast, beta-lactamases conferring low-level resistance to cefotaxime could be evolved to become high-level resistance genes in *E. coli* due to the relative paucity of chromosomal mutations conferring resistance to cefotaxime.

Two main outcomes will result from the work performed in WP4. First, the development of novel tools for metagenomic detection and characterization of minority bacterial populations present in complex samples, contaminated samples, or samples with low concentration of DNA. Second, an effective methodology for early detection and monitoring a major threat in Global Health as antibiotic resistance (Biomedicine, Food Protection, Food Safety, Environmental damage). Especially, the analysis of minority populations constitutes a major bottleneck that limits the use of current metagenomics for a broad set of applications in Biotechnology and other fields of societal interest. It is expected that the advances made in this WP will importantly:

- impact the outcomes of metagenomics applications by allowing bioprospection and characterization of minority or rare populations, significant for human, animal, food, and health environmental health.
- impact the way metagenomics is applied to solve major societal threats in Public Health, focusing in our case on a key health problem recognized as such by regulatory agencies (FMI, G8).

Due to the generic nature, this tool developed in WP4 will be (i) replicable, for verified applications, in other health institutions outside the consortium, and (ii) applicable to other health, environmental and biotechnological applications, subsequent to suitable adaptation of the detection technology (after the end of EvoTAR). Contact with RocheNimblegen has been established to improve the technology and to commercialize the design.

We foresee that the technology developed in this WP will provide measurable improvements for specific applications in Biomedicine, Agro-Food, and Environment. They include:

- *Reduction of costs* for the end users by: i) avoiding unnecessary delays in detecting threats for food safety (adoption of executive interventions); ii) improving diagnosis of Public Health threats (treatment failures, length of hospitalization stays), iii) providing biomarkers (Health) and biosensors (environment) oriented to risk assessment.
- Increasing the *quality and functionality of field and reference laboratories* by implementing standardized metagenomic procedures in routine practices (diagnosis, personalized medicine, detection of threats, management, biosurveillance, food safety, food protection, environmental risk analysis and protection).
- Increasing efficiency in development and application of *antimicrobial drugs*, and novel *decontamination and sanitation procedures* in environments, influencing global health.
- Increasing the *accuracy of predictive analysis*, assuring early containment of health threats, allowing the *definition of risks* for health related with antibiotic, metal and biocide resistance and virulence genes<sup> $2$ </sup>

The EvoTAR partners are key-opinion leaders in the fields of metagenomics and genomics, informatics and computational biology, veterinary, environmental biotechnology, infectious diseases and plasmid biology. They will draw attention to peers of EvoTAR assets, to widen the scope of the usage of this platform and to recruit future research groups for further validating of the technical approach and products. Collaboration with other research consortiums to further exploitation of the platform (environment, food production, sanitation, health, drug development) has been established. Two project proposals have been submitted, including the H2020-LEIT call which is pending of final decision, to further exploit this technology

WP6 was overall an academic work package, intending to study the dynamics of antibiotic resistance transfer within bacterial populations and to develop novel methods for this type of analysis. Therefore, results of WP6 led principally to a wide number of scientific publications, published in high-ranking journals, as their main dissemination activity. The item with the highest potential impact was the development of the bioinformatic tool called PLACNET, a method to extract and reconstruct plasmid sequences from Illumina whole-genome datasets. This is for utmost importance when studying plasmidderived antibiotic resistance whole genome sequencing (WGS) based epidemiology. WGS is more and more considered the "gold-standard" for molecular typing providing the most optimal level of resolution. However, with the use of popular short-read technologies like Illumina-based sequencing it is very difficult to reconstruct plasmids from the many assembled contigs. With the developed PLACNET tool this is now possible. The method has been made freely available for the scientific community and other potential users. We are presently working in a more use-friendly version that can be run from personal computers by using a dedicated web page.

WP7 has overall been very successful and we have generated a wealth of novel data that has been presented in publications and at various symposia, conferences, public outreach activities and seminars. With regard to impact and future use we want to point out the following broader implications of this work.

To study the between host transmission of antibiotic resistance three generic model frameworks have been developed. With the first two models it is possible to study how different diseases and antibiotic resistances spread over a network of connected hosts (i.e. persons, hospitals, farms, etc.). The framework fits well with existing data on MRSA bacteraemia cases in Scotland, thereby validating the model. These models can be used to identify hosts at risk of becoming infected and for identifying hotspots for the emergence of multidrug resistant bacteria. Furthermore, results of the models could be used to aid in the development of nationwide surveillance programs by informing policy makers where to concentrate efforts. The study on hospital level risk factors shows that, although large hospitals are important, size

<u> 1989 - Johann Stein, mars et al. 1989 - Anna ann an t-Anna ann an t-Anna ann an t-Anna ann an t-Anna ann an t-</u>

alone does not fully explain the number of bacteraemia cases in a hospital. This is of high importance for hospitals in Scotland, as this is where most of the data was collected, but also for other hospitals worldwide. Hospital systems are remarkably similar and conclusions from this study can easily be extrapolated to other hospitals or countries.

Regarding the utility of Whole Genome Sequencing of *Escherichia coli* O157 for outbreak detection and epidemiological surveillance it is demonstrated that WGS offers the potential to streamline reference laboratory processes by the use of a single diagnostic tool to generate the information required to support clinical management of cases, and Public Health investigations and interventions to control spread. It has the potential to transform the way we assess relatedness of strains and the risk of development of severe complications. However, issues relating to ease of performance and standardisation, as well as IT infrastructure and data storage need to be addressed before it is introduced routinely. Also the study on intercontinental exclusion of community-associated MRSA subtypes with distinct antibiotic resistance gene profiles demonstrates the power of whole genome sequencing for resolving longstanding questions about the origin and transmission routes of bacteria in the community setting. Through in vitro experiments inspired by our sequence analysis we also show that that direct competition is just one of a multitude of factors which determine the global distribution of bacterial strains.

The use of *in silico* pharmacokinetic-pharmacodynamic (PKPD) models based on data from *in vitro* timekill experiments can provide valuable information to guide dosing of antibiotics. Thus, the development a mechanism-based *in silico* model that can describe in vitro time-kill experiments of susceptible and resistant *E. coli* and the high correlation of killing kinetics with the experimentally determined MICs, suggest that the *in vitro* time-kill profile of a mutant strain is predictable by the MIC alone. The general applicability of this observation to other bacterial species and antibiotics is still unclear but it implies that research to identify optimal dosing regimens could be greatly simplified.

Our experimental work on fitness costs has generated some general implications of importance for understanding and predicting resistance development. Thus, almost without exception antibiotic resistance mechanisms (whether mutational or horizontally transferred) confer a fitness cost observed as a reduced growth rate/survival *in vitro* and/or *in vivo*. In general, we also observe that the fitness costs associated with various mutational resistance mechanisms are remarkably stable across different genetic contexts, bacterial species and clones, indicating that epistatic effects are weak for these particular chromosomal resistances. However, for other plasmid-borne transposons encoding antibiotic resistance only a limited number of hosts ensure their stable maintenance and here differences in fitness might explain the particular association of genetic elements (plasmids, transposons, resistance genes) with particular clones and species. One important implication is therefore that it is at present very difficult to predict the magnitude of the fitness effect of particular resistance mechanism in a particular genetic background.

In WP8, novel interventional strategies to tackle antimicrobial resistance were developed and thoroughly evaluated. Nowadays, it is well acknowledged that both new antibacterial therapies and out-of-the-box therapeutic strategies are needed to fight antimicrobial resistance. EvoTAR is contributing a lot to open new avenues for better cures and handling of patients with bacterial infections.

DAV132, developed partly in the context of EvoTAR by Da Volterra, is the first product with a clinically-demonstrated protection of intestinal microbiota from disruption during antibiotic treatments. This represents a tremendous achievement for all the teams involved in EvoTAR as it illustrates that DAV132, when administered in combination with antibiotic treatment, drastically minimizing the sideeffects on the intestinal flora; both limiting the onsets of diseases such as *Clostridium difficile* infections and curbing the emergence and spread of resistant bacteria. DAV132 is a really unique innovation and its further development up to patients has strongly been boosted by the results generated in EvoTAR. The EvoTAR project has also been beneficial for animal health as several data of importance on antibiotic resistance in animals were gathered during the work on DAV133. DAV133 is the only product in the world developed to prevent the emergence and rise of resistant bacteria in companion animals when they

receive antibiotic treatments. It has also benefited a lot from the studies performed collaboratively by EvoTAR's partners and all the EvoTAR data will support the future registration and use of the product.

Future studies will be needed to confirm the efficiency of DAV132 in protecting the microbiome from different antibiotics in a clinical context, but there is a potential that the adjunction of the product becomes a standard of care in antibiotic treatment. Health benefits of that practice could be absolutely huge, given the links between the antibiotic treatments and both the short term adverse consequences, such as *C. difficile* infections, and potential long term consequences, linked to loss of gut microbial richness, associated to the risk of pathologies associated to obesity and the metabolic syndrome, such as type 2 diabetes, hepatic and cardiovascular complications and certain cancers.

Work done in WP8 by Instituto Biomar and its partners in Universidad de Cantabria has allowed characterizing of conjugation inhibitors, and the evaluation and demonstration of their potential to prevent the transfer of antibiotic resistance genes through different conjugation systems. These results could lead to the development of products based on the inhibition of conjugation to be administered concomitantly with antibiotic applications in different settings. This work has resulted in scientific publications, and if the challenges remaining in the route to create commercial products based on these properties are overcome, the Evotar project will have been essential in the development of these products, that will have a huge social and economic impact.

# **Project web-site: www.evotar.eu**

# **Coordinator:**

Rob J. L. Willems Universitair Medisch Centrum Utrecht, Department of Medical Microbiology, Heidelberglaan 100, CX Utrecht, 3584 CX, the Netherlands r.willems@umcutrecht.nl

# **Partners:**

Teresa Coque / Fernando Baquero Servicio Madrileno de Salud, Microbiology Department, Ctra.Colmenar Viejo km 9.100, Madrid, 28034, Spain baquero@bitmailer.net

José Luis Martínez Menéndez Agencia Estatal Consejo Superior De Investigaciones Cientificas, CNB - Microbial Biotechnology Department, Darwin 3, Madrid, 28049, Spain jlmtnez@cnb.csic.es

Dan Andersson Uppsala Universitet, Department of Medical Biochemistry and Microbiology, Husargatan 3, Uppsala, 75123, Sweden Dan.Andersson@imbim.uu.se

Morten Sommer Danmarks Tekniske Universitet, Department of Systems Biology, Anker Engelundsvej 1, Building 101ª, Kongens Lyngby, 2800, Denmark
## msom@bio.dtu.dk

Hauke Smidt Wageningen Universiteit, Dept. Agrotechnology and Food Sciences, Laboratory of Microbiology, Dreijenplein 10, Wageningen, 6703BH, Netherlands Hauke.Smidt@wur.nl

Luca Guardabassi Københavns Universitet, Dep. Veterinary Disease Biology, Faculty of Life Sciences, Sitgbøjlen 4, Frederiksberg, 1870, Denmark lg@life.ku.dk

Antoine Andremont Assistance Publique – Hospitaux de Paris, 3, Avenue Victoria, Paris Cedex 75004, France antoine.andremont@bch.aphp.fr

Florence Sejourne Da Volterra SAS, Rue de Charonne 172, Paris, 75011, France florence.sejourne@davolterra.com

Fernando de la Cruz Universidad De Cantabria, Instituto de Biomedicina y Biotecnología de Cantabria-Intergenomics Group, Facultad de Medicina. Av. Cardenal Herrera Oria 2, Santander, 39011, Spain fernando.cruz@unican.es

Didier Mazel Institut Pasteur, Dep. Genomes and Genetics/Bacterial Genome Plasticity, Rue Du Docteur Roux 25-28, Paris Cedex 15, 75724, France mazel@pasteur.fr

Antonio Fernandez Medarde Instituto Biomar SA, Calle D Parque Tecnologico De Leon Parcela M 10.4 Armunia, Leon, 24009, Spain afernandez@institutobiomar.com

Bruno Gonzalez-Zorn Universidad Complutense De Madrid, Departamento de Sanidad Animal, Facultad de Veterinaria, Avda Puerta de Hierro s/n, Madrid, 28040, Spain bgzorn@vet.ucm.es

S. Dusko Ehrlich

Institut National De La Recherche Agronomique, MICA Division – INRA, Centre de Recherche de Jouy en Josas, Jouy en Josas, 78352 cedex, France dusko.ehrlich@jouy.inra.fr

Mark Woolhouse The University of Edinburgh, Centre for Infectious Diseases, Ashworth Laboratories, The King's Buildings, West Mains Road, Edinburgh, EH93JT, United Kingdom Mark.Woolhouse@ed.ac.uk

Gian Maria Rossolini Universita' Degli Studi Di Siena, Unit of Microbiology and Virology, Via Banchi Di Sotto 55, Siena, 53100, Italy rossolini@unisi.it

Colin Ingham H. Kruyt building O-804/805, Padualaan 8, 3584 CH Utrecht, The Netherlands colin.ingham@wur.nl , info@microdish.nl



## Section A (public) **Section A (public)**









































































































































































































## Section B (Confidential or public: confidential information to be marked clearly) **Section B (Confidential or public: confidential information to be marked clearly)**

## **Part B2**

