Aix-Marseille Université

Faculté de Médecine et de Pharmacie de Marseille

Ecole Doctorale des Sciences de la Vie de la Sante (EDSVS)

THÈSE DE DOCTORAT

Présentée et soutenue le 15 Décembre 2016

Par

Mme **TERESA CIMMINO**

En vue de l'obtention du grade de **DOCTEUR** de **l'Université Aix-Marseille**

Spécialité: Pathologie Humaine-Maladies Infectieuses

Whole genome sequencing to decipher the resistome of clinical multidrug-resistant bacteria

Membres du Jury

Président du Jury	Professeur Pierre-Edouard Fournier
Rapporteur	Docteur Marie Kempf
Rapporteur	Professeur Marie-Hélène Marchandin
Directeur de thèse	Professeur Jean-Marc Rolain

Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes

UM63 – CNRS 7278 – IRD 198 - INSERM U1095

IHU Méditerranée Infection

TABLE OF CONTENTS

- **CHAPTER IV (ANNEX): Collaborative working of description of new bacterial species**.

  - **Article 6:** Noncontiguous finished genome sequence and description of *Bacillus andreraoultii* strain SIT1T sp. nov. S. I. Traore, T. Cimmino, J.-C. Lagier, S. Khelaifia, S. Brah, C. Michelle, A. Caputo, B. A. Diallo, P.-E. Fournier, D. Raoult and J. M. Rolain. **New Microbes and New Infections.**

  - **Article 7**: Noncontiguous finished genome sequence and description of *Paenibacillus antibioticophila* sp. nov. GD11T, the type strain of *Paenibacillus antibioticophila*. G. Dubourg, T. Cimmino, S. a. Senkar, J.-C. Lagier, C. Robert, C. Flaudrops, P. Brouqui, D. Raoult, P.-E. Fournier and J.-M. Rolain. **New Microbes and New Infections.**

  - **Article 8**: Noncontiguous finished genome sequence and description of *Enterococcus massiliensis* sp. nov. S. Le Page, T. Cimmino, A. Togo, M. Million, C. Michelle, S. Khelaifia, J.-C. Lagier, D. Raoult, and J.-M. Rolain. **New Microbes and New Infections.**

  - **Article 9**: Non contiguous-finished genome sequence and description of *Murdochiella massiliensis* strain SIT12 sp. nov. Elisa Vicino, Sory Ibrahima Traore, Teresa Cimmino, Grégory Dubourg, Noemie Labas, Claudia Andrieu, Fabrizio Di Pinto, Cheikh Sokhna, Aldiouma Diallo, Didier Raoult, Jean Marc Rolain. **New Microbes and New Infections.**

  - **Article 10**: Genome sequence and description of *Mobilicoccus massiliensis* sp. nov. isolated from the stool of a Nigerien male suffering from a severe form of acute malnutrition ‘‘kwashiorkor**’’**. Najla Mathlouthi Sory Ibrahima Traore, Teresa Cimmino, Saber Khelaifia, Thi Tien Nguyen, Frederic Cadoret, Carine Couderc, Didier Raoult and Jean Marc Rolain. **New Microbes and New Infections.**

- **CHAPTER V: Conclusion and perspectives.**
- **Acknowledgements**

AVANT PROPOS

Le format de présentation de cette thèse correspond à une recommandation de la spécialité Maladies Infectieuses et Microbiologie, à l'intérieur du Master des Sciences de la Vie et de la Santé qui dépend de l4 école Doctorale des Sciences de la Vie de Marseille.
Le candidat est amené à respecter des règles qui lui sont imposées et qui comportent un format de thèse utilise dans le Nord de l'Europe et qui permet un meilleur rangement que les thèses traditionnelles. Par ailleurs, la partie introduction et bibliographie est remplacée par une revue envoyée dans un journal afin de permettre une évaluation extérieure de la qualité de la revue et de permettre à l'étudiant de commencer le plus tôt possible une bibliographie exhaustive sur le domaine de cette thèse.
Par ailleurs, la thèse est présentée sur article publié, accepté ou soumis d'un bref commentaire donnant le sens général du travail. Cette forme de présentation a paru plus en adéquation avec les exigences de la compétition internationale et permet de se concentrer sur des travaux qui bénéficieront d'une diffusion internationale.

Professeur Didier RAOULT

RESUME

L'émergence des nouvelles technologies de séquençages (NTS) à révolutionné le domaine médical particulièrement dans l'amélioration des pratiques médicales telles que l'optimisation des outils de diagnostic, la prise en charge thérapeutique, et le traitement des patients. Le séquençage des génomes bactériens d'intérêt clinique a permis aujourd'hui, non seulement l'identification d'un pathogène dans un échantillon donné, mais aussi l'étude de sa pathogénicité, ainsi que la détection des facteurs de virulence (virulome) et des gènes de résistance aux antibiotiques (résistome). Dans le cadre de l'étude du résistome des Bactéries MultiRésistantes (BMR), l'utilisation des NTS permet aujourd'hui d'analyser de manière exhaustive et de comprendre les différents mécanismes de résistance, leurs répertoires génétiques, ainsi que leurs mécanismes de dissémination au niveau mondial.

C'est dans cette optique que ma thèse s'inscrit avec comme objectif: « Le séquençage de génomes de bactéries multi résistantes d'intérêt clinique pour définir leur résistome ». Ainsi, au cours de ces trois années de thèse, nous avons pu réaliser :

1. Une revue de la littérature sur l'utilisation des nouveaux outils de diagnostic contemporains et leurs capacités a détecter des épidémies causées par les BMR.

2. La détermination et l'analyse du resistome de génomes d'isolats cliniques multirésistants telles que *Shewanella algae*, une bactérie de l'environnement, isolée d'un prélèvement respiratoire chez un patient hospitalisé atteint de pneumonie et *Chryseobacterium indologenes*, isolée chez une patient atteint de mucoviscidose. Au cours de cette analyse, nous avons pu montrer que les bactéries de l'environnement telle que *S. algae* peuvent représenter un réservoir de gènes de résistance aux antibiotiques. L'analyse exhaustive de ces bactéries a également montré la capacité de ces dernières à s'adapter à leurs

écosystèmes notamment par l'acquisition de nouveaux éléments génétiques par transfert latéral de gènes.

Par ailleurs, la détection de gènes impliques dans la synthèse de Non-ribosomale Polypeptides Synthases (NRPS) et de Polypeptides Synthases (PKS) pourrait avoir un rôle dans leur capacité a survivre dans des milieux hostiles tel que le tractus respiratoire de patients mucoviscidosiques ou leur présence chez des patients ayant subi de multiples lignes d'antibiothérapies.

3. Dans ce travail, grâce à l'utilisation des NTS sur des nouvelles espèces bactériennes isolées du microbiote humain, nous avons pu réaliser une analyse standardisée "*in silico*" afin de déterminer le résistome de ces bactéries ainsi que la présence de métabolites secondaires associés aux bactériocines et aux NRPS/PKS. L'application des NTS pour le séquençage du génome complet de nouvelles espèces bactériennes isolées dans le microbiote humain, nous a permis de développer une plateforme capable d'analyser ces nouvelles espèces en moins de 48 h. Ce travail permet d'avoir une meilleure compréhension de la biodiversité des bactéries isolées au sein du microbiote humain.

SUMMARY

The emergence of new sequencing technologies (NTS) have revolutionized the medical field improving medical practices such as the optimization of diagnostic tools, therapeutic care and treatment of patients. Whole genome sequencing allows the identification of bacteria from clinical sample and the study of its pathogenicity as well as the detection of virulence factors (virulome) or antibiotic resistance genes (resistome).The use of NTS allows to analyze and to decipher the study of resistome of Multi-Drug Resistant bacteria (MDR), understanding the different resistance mechanisms, genetic directories and their dissemination mechanisms at global level. It is in this context that my thesis project enrolls with the main objective using:"Whole genome sequencing to decipher the resistome of clinical multidrug resistant bacteria". During the three years of thesis, we have achieved:

1. A literature review on the use of new contemporary diagnostic tools and capabilities in detecting outbreaks in infectious diseases caused by MDR.

2. The identification and the analysis of resistome of multidrug resistant bacteria from clinical isolates such as *Shewanella algae*, normally marine environmental, in our case clinical strain isolated from the bronchoalveolar lavage (BAL) of a hospitalized patient with pneumonia and *Chryseobacterium indologenes*, isolated from a patient cystic fibrosis. In this analysis, we can show that environmental bacteria such as *S. algae* can be a reservoir of antibiotic resistance genes. The exhaustive analysis of these bacteria showed their ability to adapt to their ecosystems including the acquisition of new genetic elements by lateral gene transfer.

Furthermore, the detection of genes involved in the synthesis of nonribosomal peptide synthetase and polyketide synthases may have a role in their ability to survive in hostile environments such as the respiratory tract of CF patients or their presence in patients having suffered multiple of antibiotic.

3. In this work, through the use of the NTS on new bacterial species isolated from human microbiome, we have achieved a standardized analysis "*in silico*" to determine the resistome of these bacteria and the presence of secondary metabolites associated bacteriocins and the NRPS / PKS. The application of the NTS for sequencing of bacterial genome of new bacterial species isolated in the human microbiome, allowed us to develop a platform capable of analyzing these new species within 48 hours. This work provides a better understanding of the biodiversity of bacteria isolated in the human microbiome.

INTRODUCTION

The discovery of natural and synthetic compound able to eradicate infectious diseases, defined incurable, has revolutionized modern medicine. By definition, antibiotics (from the Greek αντι,"against" and βιος, "life") are substances produced by microorganisms and they have antagonistic effects on the growth of other microorganisms [1]. In this context, intense scientific research led to the various market drugs having antimicrobial activity. Following the success of penicillin in controlling bacterial infection among the soldiers during the Second World War, emergence of penicillin-resistant strains was evidenced in the 1940s. In the late 1960s, the unprecedented success of early antibiotic therapies led US Surgeon General William H. Stewart to make the famous declaration: "it is time to close the book on infectious diseases and declare the war against pestilence won"[2]. The euphoria did not last long and the magic bullets started losing the efficacy because of the steady emergence of antibiotic resistant pathogens simultaneously with their widespread use. By 1960, it assumed the shape of a pandemic problem. New β-lactam antibiotics were introduced into clinical practices to restrain the problem. Simultaneously, bacterial strains resistant to them came into being, a phenomenon dubbed as β-lactamase cycle. The first case of MRSA was identified in the UK in 1961 [3].

During the last 40 years, the dissemination of multi-drug resistant bacteria (MDRB) has become a major public health concern worldwide because of the increase of infections caused by MDRB, the difficulty in treating them and the expenditures in patient care [4]. The massive use of antibiotics in human medicine and in animals, combined with the possibility of the emergence at any time of a new bacterial clone with high fitness and virulence, even if it is not linked to selection pressure, may also create a stochastic event associated with a MDRB; Alice's living croquet theory [5]. The bacterial adaptation contributes to the dissemination of new clones and/or new genes of interest such as resistance genes. Antibiotic resistance is the result of different and several factors, that are unstoppable because of the high diversity and capacities of microbes to adapt to their environments [6]. The development of resistance is inevitable, it is impossible to prevent the emergence of resistance, rare, random and usually transient, if it does not provide a selective advantage. A winning strategy is knowledge of the «enemy». The rapid detection of resistance mechanisms has an important epidemiological role. This information is critical for clinicians so that they choose the best antimicrobial treatment that will reduce the burden of hospital-acquired infection caused by MDR pathogens. The knowledge of the mechanisms of bacterial resistance and the

development of new drugs with antimicrobial activity are essential to limit the damage and keep open a chance of cure.

As proposed by Gerard D. Wright, the resistome is defined as the collection of all antibiotic resistance genes. It includes resistance elements found in both pathogenic bacteria and antibiotic-producing bacteria, and cryptic resistance genes (which are not necessarily expressed) that are present in bacterial chromosomes [6]. It is in this context that my thesis project enrolls with the main objective of using the whole genome sequencing to analysis and to decipher the resistome of clinical isolate of multidrug-resistant to uncover the genomic origins of epidemiologic effect. Through the advances of high-throughput sequencing technologies and the open source of bioinformatic tools, we can identify pathogens and their resistance profiles very rapidly. Bioinformatics tools provide better understanding of the molecular basis of antimicrobial resistance.

Finally, this manuscript presents the *cursus* of my thesis articulated into four parts:

First chapter: The first part is dedicated to a literature review published in the Expert Review of Molecular Diagnostics (Impact factor 3.3)(**Article 1**). The goal was to summarize and to explain the contemporary challenges and opportunities in the 21st century to manage the problem of multi-drug resistant bacteria (MDRB). The need for faster provision of care, active surveillance of circulating bacteria, rapid detection of an abnormal event (new or emerging phenotype), development of rapid tests for the detection of new cases (phenotypic tests and real-time PCR assays based on whole genome sequence analysis of the clone) to prevent the outbreak, and the need for rapid action to avoid dissemination of the infections disease caused by MDRB.

Second chapter: I present in this chapter the studies performed using whole genome analysis sequencing of MDRB *Shewanella algae* (**Article 2**) and *Chryseobacterium indologenes* (**Article 3**) isolated from clinical samples. The choice to study these bacteria has been developed respectively following an imipenem and colistin resistance during patient hospitalisation. Particularly, *Shewanella algae* is widespread in the environment and is part of the marine microflora. *Shewanellae spp*. are increasingly being implicated as human pathogens in people exposed to marine niches activities containing *Shewanellae.* In our case, this clinical strain was isolated from the bronchoalveolar lavage (BAL) of a patient with aspiration pneumonia at the Timone Hospital, in Marseille, France after plunging into the Mediterranean Sea. Moreover, *Chryseobacterium indologenes* is an environmental organism and usually associated with nosocomial infections. This time, the bacterium was isolated from

a respiratory sample of cystic fibrosis (CF) patient hospitalized in the Timone Hospital in Marseille.

Third chapter: the culturomics, a new strategy to study the human microbiome allowed the identification of many new bacterial species. This new approach has revolutionized the bacterial culture. Using the combination of 16S rRNA bacterial, Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS) and Whole genome sequencing (WGS). The rapidly increasing number of available genome sequences has proposed the creation of new bacterial species based on this strategy which we call taxonogenomics. Taxonogenomics consists of a polyphasic approach, incorporating both phenotypic and genotypic data, to describe a new bacterial species. First, we consider bacterial strains that exhibit a 16S rRNA sequence identity of 98.7% compared with the phylogenetically most closely related species with standing in nomenclature, as previously recommended, as belonging to a putative new species, provided that this identity value is in the range of identities observed among validly published species within the same genus. Next, we study the primary phenotypic characteristics of the bacterial isolate, such as habitat, Gram staining, electron microscopy, primary culture, and metabolic characteristics; proteic spectra obtained by MALDI-TOF MS; and genomic characteristics (genome size; G C content; percentage of coding sequences; gene content; gene distribution in COG categories). Type strains are also deposited into two international culture collections, including our collection (Collection de Souche de l'Unité des Rickettsies [CSUR] [7]. My research work was performed on new bacterial species focusing on genomic analysis by standardization of "*in silico"* methods to decipher the resistome and novel secondary metabolites such as bacteriocin and polyketide synthases/ non-ribosomal peptide synthetases (PKS/NRPS). This metagenomics approach has resulted in published articles on description on new species (**Article 4-5**).

Fourth chapter (annex): This chapter is devoted to the collaborative research works. I was involved in the study of antibiotic resistance (resistome) and secondary metabolites focusing on bacteriocin PKS/NRPS in genome of new bacterial species. The application of WGS in this study revealed the resistome and the ability to produce a secondary metabolite of commensal bacteria resides on or within a number of tissues including skin and human gut. This research project shed light the different features of bacteria living human body showing the relationship between health and non-health human and changes in the human microbiome. The collaborative research works has resulted in analyzing of more than 10 genome of new

bacterial species with 4 published article (**Article 6-7-8-9**) two submitted article (**Article 5-10**) and 11 articles in progress.

References

1. Sengupta S, Chattopadhyay MK, Grossart H-P. The multifaceted roles of antibiotics and antibiotic resistance in nature. *Front. Microbiol.* 4, 47 (2013).

2. Spellberg B, Guidos R, Gilbert D, *et al.* The epidemic of antibiotic-resistant infections: a call to action for the medical community from the Infectious Diseases Society of America. *Clin. Infect. Dis.* 46(2), 155–64 (2008).

3. Johnson AP. Methicillin-resistant *Staphylococcus aureus*: the European landscape. *J. Antimicrob. Chemother.* 66 Suppl 4, iv43–iv48 (2011).

4. Chang H-H, Cohen T, Grad YH, Hanage WP, O'Brien TF, Lipsitch M. Origin and proliferation of multiple-drug resistance in bacterial pathogens. *Microbiol. Mol. Biol. Rev.* 79(1), 101–16 (2015).

5. Raoult D. Alice's living croquet theory. *Int. J. Antimicrob. Agents*. 47(4), 249 (2016).

6. Wright GD. The antibiotic resistome: the nexus of chemical and genetic diversity. *Nat. Rev. Microbiol.* 5(3), 175–86 (2007).

7. Lagier J-C, Hugon P, Khelaifia S, Fournier P-E, La Scola B, Raoult D. The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. *Clin. Microbiol. Rev.* 28(1), 237–64 (2015).

Second chapter

The emergence of next-generation sequencing (NGS) technologies, have revolutionized microbiology. Whole genome sequencing (WGS) is an emerging technique through that molecular characterization of isolates on many levels, without the need for a priori selection of targets as is required for several standard techniques such as PCR. One of the primary investigation approach in microbial research is the use of genomics to characterize an organism, including identification of the genetic elements that may result in pathogenicity, survival, or antimicrobial resistance. As with human genetics, microbial genomics has the capacity to interrogate organisms for key genetic markers that may influence treatment and prognosis of infections [1]. WGS offers great promises, as the repertoire of resistance genes ('the resistome'), the chromosomal background of strains and the plasmid sequences can be deduced from the same data [2].

In this context the works on *Shewanella algae* and *Chryseobacterium indologenes* were performed. These multidrug-resistant bacteria (MDRB) were isolated from hospitalized patient undergoing antibiotic therapy. In critical care units, there is extensive antimicrobial use, which imposes a selection pressure and promotes the emergence of MDRB. The aim of these studies is to improve the knowledge of bacterial resistome to predict the optimal antibiotic treatment and to decrease the spread of MDR nosocomial bacteria. In 2011, a 39 year-old man accidentally plunged into the Mediterranean Sea and was later admitted to the La Timone Hospital in Marseille. He suffered from a broken neck and aspiration pneumonia due to lung aspiration with seawater. *S. algae* was later isolated in pure culture from his bronchoalveolar lavage (BAL). The Antibiotic susceptibility testing of these bacteria revealed the resistance to imipenem, ticarcillin/clavulanic acid and piperacillin. In La Timone hospital, the first identification by MALDI TOF MS revealed the *S. putrefaciens* strain but using a more detailed analysis of *rpoB* the analysis confirmed *S. algae* strain. *Shewanellae* spp. are common bacteria members of a complex community in aquatic and sedimentary environments that are chemically stratified on a permanent or seasonal basis, the limited number of human clinical specimens examined report that *S. putrefaciens* and *S. algae* are most commonly isolated. Due to the infrequent update of these two *Shewanella* species in the databases of automated identification systems, the clinical laboratory may encounter difficulty in reliably differentiating between *S. algae* and *S. putrefaciens*. *S. algae* has been mainly reported from human clinical sample, while *S. putrefaciens* was time isolated from non-human sample such as freshwater fish and water. The genome analysis of *S. algae* allowed providing a detailed

repertoire of resistance genes including the presence of members of class C β-lactamases *ampC* never identified before in *Shewanella* spp., the mapping gene of Ambler class D β-lactamases-encoding bla-$_{oxa55}$ gene. At the same time, the potential role of this bacterium as reservoir and vector of antimicrobial resistance was confirmed although the presence of *qnr* encoding gene already known as the source of plasmid-mediated QnrA determinants [3]. *Shewanella algae*, already known as an environmental genus reveals the presence of genes associated with resistance to antiseptics and heavy metals that could be used as alternative electron acceptors enabling it to grow under extreme and varied conditions. The respiratory diversity of these organisms is one of their greatest assets in terms of survival in the environment and in human infections. Similarly, complete flagellar system localized in a single locus, on the chromosome and probably derived from ancient hemolysins genes and genes responsible for biofilm formation that allow to the bacterium to adapt and survive in harsh environments such as seawater and sewage and also enabling it to cause severe human infections (**Article 2**). The fascinating physiology of these organisms had led to several biotechnological uses and the same *S. algae* strain will be used to test new antimicrobial activity thanks the presence of bacteriocin with uncharacterized antibacterial activity. Preliminary tests show that *S. algae* can inhibit the growth of *Candida* spp*, K. pneumoniae and Enterobacter cloacae.* Comparative studies (data not published) carried out on other *Shewanella* have revealed that the presence of bacteriocin operon (a marinocin-like, broad-spectrum antibacterial protein) found only in our strain and was absent in all *Shewanellae spp.* analyzed.

The same approach was used to study *Chryseobacterium indologenes* MARS15; this bacterium is resistant to several antimicrobial drugs, particularly colistin. The prevalence of *Chryseobacterium* infection in cystic fibrosis patients have been studied previously [4]. In 2013, *C. indologenes* was isolated from the sputum sample of a 15 years old CF girl, admitted to the La Timone Hospital in Marseille, regularly treated with aerosolized colistin for chronic *Pseudomonas aeruginosa* exacerbations **(Article3)**. After the identification by MALDI TOF MS and the E-test was performed and show high resistance to colistin (256 µg/ml), amoxicillin (256 µg/ml), and imipenem (32 µg/ml). Deciphering the resistome of this bacterium, we identify a reservoir of diverse β-lactamases of the Ambler class A β-lactamase encoding bla$_{CIA}$ and of the Ambler class B β-lactamase bla$_{IND-2}$, four MBL (metallo-β-lactamase) metallohydrolases, a chloramphenicol acetyltransferase encoding ACT, monoxygenase responsible for tetracycline resistance and the multidrug efflux pump AcrB, known to be involved in drug resistance. Apart from deciphering the resistome of this atypical

bacterium, we identified three specific features in the *C. indologenes* MARS15, genome. First, the presence of Polyketide Synthases (PKS) and Nonribosomal peptide synthetases that are a class of organic compounds usually produced of a secondary metabolite. Second, an urease operon, was identified. Urease enzyme is a widely expressed virulence factor. This operon could allow bacterial survival in the acid microenvironment of the inflamed human respiratory tracts of patients with COPD or CF as in our case. In addition, the release of the strongly alkaline agent ammonia and its derivatives by ureolysis could damage host tissues, which may exert its greatest impact in pathogenesis [5]. Third, the yellowish pigmentation in *C. indologenes* could play a virulence protective role by allowing a given microbe to evade host immune killing or by provoking inflammatory damage to cells and tissue. Furthemore, *C. indologenes* MARS15 intact phage absence in the other strain may confer a competitive advantage in chronic lung infection.

References

1. Kwong JC, Mccallum N, Sintchenko V, Howden BP. Whole genome sequencing in clinical and public health microbiology. .

2. Bootsma HJ, Schouls LM. Next-generation sequencing of carbapenem-resistant Gram-negative microorganisms: a key tool for surveillance and infection control. *Future Microbiol.* 10(3), 299–302 (2015).

3. Nordmann P, Poirel L. Emergence of plasmid-mediated resistance to quinolones in Enterobacteriaceae. *J. Antimicrob. Chemother.* [Internet]. 56(3), 463–9 (2005). Available from: http://www.ncbi.nlm.nih.gov/pubmed/16020539.

4. Lambiase A, Del Pezzo M, Raia V, Sepe A, Ferri P, Rossano F. Chryseobacterium respiratory tract infections in patients with cystic fibrosis. *J. Infect.* [Internet]. 55(6), 518–23 (2007). Available from: http://www.ncbi.nlm.nih.gov/pubmed/17889369.

5. Zhang L, Patel M, Xie J, Davis GS, Marrs CF, Gilsdorf JR. Urease operon and urease activity in commensal and disease-causing nontypeable Haemophilus influenzae. *J. Clin. Microbiol.* 51(2), 653–5 (2013).

CONCLUSION

It is important for hospitals to improve the processes of care known to impact nosocomial infection rates. Determining the antimicrobial patterns of the disease causing by organisms will enable health institutions to restrict the use of antimicrobials and take active measures in preventing the spread of drug resistance in hospitals. The rapid detection of resistance mechanisms has also an important epidemiological role.

The analysis using WGS has the potential to decipher the mechanism of dissemination of a MDR and possible outbreak in a short time. Colonization of MDRB such as *Chryseobacterium indologenes* could be not relevant in healthy individual but, its presence in immunocompromised patient is more dangerous. At the same time, the bacteria not eradicate can transmitted from one patient to another and to contaminate all the area in the hospital. Moreover, the study of resistome allow to select the optimal treatment for the specific microorganism and to improve the patient's life.

The knowledge of the “enemy” is a winning strategy. The bacterial genomes carrying little, if any, non-functional DNA, that can challenge in two ways. First, they often had large expansions of repetitive, selfish DNA in the form of insertion sequences. Second, they often had large numbers of pseudogenes—genes inactivated by point mutation or disruption, sometimes involved in pathogenicity or host interactions.

We can conclude that one common factor for these two MDRB is that they are a reservoir of antibiotic resistant genes. *S.algae*, represents not only a potential reservoir but also a vector of antimicrobial resistance mechanisms in hospital settings and the environment, with transmission occurring in both directions. Our studies show its characteristic of reservoir for oxacillinase genes. Similarly, *C. indologenes* is the reservoir of metallo-beta-lactamases (MBL), which potentially can spread to gram-negative bacteria of greater clinical significance.

Genome approaches provided information on repertoire of resistance and virulence determinants and host-interaction factors, allowing the proposal of hypotheses to explain the evolution of many of these traits.

The judicious use of antibiotics by health workers and efforts to control the use of antibiotics remains necessary to help to limit the increasing rates of multi-drug resistance in pathogens.

Third Chapther

The exhaustive description of human microbiota and their relationship with health and disease are major challenges in the twenty-first century [1]. Molecular methods, particularly those based on sequencing of the 16S rRNA gene, have enabled a striking increase in the number of identified environmental and clinical bacterial species that were previously unidentifiable with phenotypic methods [2]. New notion termed taxonogenomics to describe novel bacterial species is based on a polyphasic approach that included phenotypic as well as genomic criteria (genome characteristics as well as genomic sequence similarity.

Microbial culturomics, a recent concept based on a use of several culture conditions with identification by MALDI-TOF followed by the genome sequencing of the new species cultured had allowed a complementarity with metagenomics [3]. The complement of microorganisms that live on and within us, our microbiome, and its role in health and disease has become a central focus of current research [4]. Thanks to genomic analysis we can investigate functional role of the resistome on the carriage and transmission of AROs, provide a comprehensive picture of the resistance "potential" of a community. An advantage of metagenomics is that as new tools are created and new discoveries made the sequencing data can be utilized in ways previously unknown.

During the studies on new bacterial species isolated from human microbiome I analyzed more than ten genomes of different bacterial genomes among positive Gram and negative Gram, isolated from healthy and disease individual. Use of standardized bioinformatics tools to decipher the resistome, to identify the bacteriocin and secondary metabolites focusing on NRPS (non-ribosomal peptide synthetases) and PKS (polyketide synthases) have enabled to retrieve the results in less than 48 h. This work revealed the reservoir of ARG (Antibiotic Resistance Gene) in nonpathogenic bacteria such as in *Pelistega massiliensis* (data not yes published), human commensal such as *Actinomyces polynesiensis* str. MS2 (**Article 3**), *Bacillus testis* SIT10 (**Article 4**), emerging pathogens *Chryseobacterium timonense* G972 (**Article5**). Moreover, the presence of colicin, belonging to the family of bacteriocins was found in *Bacillus testis* SIT10 (**Article 3**).

References

1. Lagier J-C, Armougom F, Million M, *et al.* Microbial culturomics: paradigm shift in the human gut microbiome study. *Clin. Microbiol. Infect.* 18(12), 1185–93 (2012).

2. Perrine Hugon, Jean-Charles Dufour, Philippe Colson, Pierre-Edouard Fournier, Kankoe Sallah DR. A comprehensive repertoire of prokaryotic species identified in human beings. - PubMed - NCBI. *Lancet Infect Dis.* (2015).

3. Jean-Christophe Lagier, Matthieu Million, Perrine Hugon FA and DR. Human gut microbiota: repertoire and variations. - PubMed - NCBI. *Front Cell Infect Microbiol.* (2012).

4. Mulcahy-O'Grady H, Workentine ML. The Challenge and Potential of Metagenomics in the Clinic. *Front. Immunol.* 7, 29 (2016).

Conclusion

The bacteria that can be grown in the laboratory are only a small fraction of the total diversity that exists in nature. Approximately only 1% of bacteria on earth can be readily cultivated in vitro. However, one of disadvantage of these cultured-based techniques is that they are limited to screening for antibiotic susceptibility profile. For example, the antimicrobial susceptibility test (AST) of obligate anaerobic microorganism is not standardardized and sometimes not possible to reproduce.WGS approach can analyze the resistome of bacteria in depth without the need of culture. During our work on strict anaerobic microorganisms, it was possible through *in silico* analysis to retrieve the genotypic results before obtaining the *in vitro* results of AST. WGS methods may be useful for slow-growing organisms, organisms that are unable to be cultured, or where phenotypic susceptibility testing is unreliable. The WGS analysis in culturomics approach allowed to improve the knowledge of human gut, how microbiome human changes relating to various countries, food traditions and health conditions. The information retrieved from WGS analysis of new bacterial strains revealed how 'friendly' gut bacteria interact with the body to influence nutrition and disease. Moreover, sequencing approach would be used as a single comprehensive screening test for potential pathogens, both known and novel, as well as to assess the state of an individual’s microbiome.

Fourth chapter

In this section, we present to the collaborative research works dealing with analyzing of new bacterial species. The study allowed to discovery reservoir of ARG (Antibiotic Resistance Gene) *Pelistega massiliensis* (data not yes published), human commensal such as *Bacillus andrearaoultii* SIT-1 (**Article 6**), *Enterococcus massiliensis* AM1 (**Article 8)** and emerging pathogens *Chryseobacterium timonense* G972 (**Article5**). Moreover, the presence of colicin, belonging to the family of bacteriocins was found in *Bacillus andrearaoultii* SIT-1 and secondary metabolite (NRPS and/or PKS) that remains silent or inactive under normal laboratory conditions as, *Chryseobacterium timoniense* G792, *Pelistega massiliensis* MC2, (data not yet published) *Paenibacillus reamassiliensis* GD6 (data not yet published), *Paenibacillus touaregensis* Marseille-P2472T (data not yet published), *Paenibacillus rubiinfantis* (data not yet published) and *Paenibacillus senegalomassiliensis* (data not yet published), Clostridium niameyense MT (data not yet published), *Paenibacillus numidis* GM2 (data not yet published).

Conclusion and Perspectives

In this thesis, we decipher the resistome and analyze the genome of clinical multidrug resistant bacterial species. We discovered new resistance genes in emerging multidrug-resistant bacteria such as *Shewanella algae* and *Chryseobacterium indologenes* and new mechanisms that might be responsible of synthesis of new drugs or uses in the field of biotechnology. In routine clinical practice and in the future, these results can be used to improve the knowledge of bacteria to understand the network of Antibiotic Resistance Gene from the environment to human for controlling and predicting the emergence of antibiotic resistance in the clinic and treat the patient using the antibiotics. The knowledge could transform the diagnosis, treatment and predictability of bacterial infections. Importantly, genotypes can inform not only on the current drug susceptibility of a pathogen but also on its future potential to evolve resistance and spread. For example, sequencing of bacterial genome could determine whether a drug susceptible strain carries precursors to resistance genes (which are termed proto-resistance genes), such as drug-degrading enzymes or efflux pumps, that might be mutated to increase expression or to strengthen activity. Our knowledge of the human resistome remains low and future studies should be focused on discovery of these new antibiotic resistance genes for a better understanding of the magnitude of this phenomenon among environmental reservoirs and its potential impact in human pathogens [1].

In this context, the development of the resistome study of new bacterial species can reveal the presence of reservoir of ARGs where the bacterium is part of the commensal bacterial flora of a healthy people. The knowledge of these proto-resistance genes will help the personalized antibiotic treatment of the patients, preventing the spread and buildup of resistance antibiotic gene. Based on the data and results of my Phd thesis, I draw a useful conclusion that whole genome sequencing of multidrug resistant isolates has allowed to decipher the genetic determinants and allow to unravel intrinsic mechanisms by which bacteria are equipped and they use only when it is necessary. At present, whole genome sequencing can detect pathogen resistance very rapidly, and may help in responding efficiently to clinical questions. Despite the fact that NGS studies have demonstrated high concordance between *in silico* predicted and phenotypic antimicrobial susceptibility, we add a note of caution, since the genome sequence does not as yet allow accurate prediction of the potentially conditional expression of particular genes, or their expression level. Pathogens are not static: they move around the world, they adapt to the environment, their lifestyles change,

and so the need to create systematic computational methods, a unique, easily accessible and exhaustive antibiotic resistance genes database and a common web-based database network will be necessary in the future in order to exchange scientific information for global surveillance. Genome sequencing will become a mainstay in analyses of infectious agents for both routine (e.g. diagnosis) and high-resolution (e.g. outbreak tracking) purposes. Moreover, it will involve on detailed characterization of bacterial genomes, confirming putative virulence factors and facilitating an understanding of the disease process. The current and future study on the bacterial resistome will be used to improve the knowledge of the characteristics of bacteria to be able to manage the appropriate treatment for patients.

Metagenomics-based approaches might become front line diagnostic test for infectious disease in the Public health setting; When encounter an new or complex infectious disease, multiple conventional diagnostic test are often used, potentially leading to unnecessary cost and delays in diagnosis. Moreover, the decrease in price and increase in speed and simplicity of sequencing, it will become possible to personalize the therapies for individual treatment of infectious diseases and to improve food protection strategies. This has proved very effective in the surveillance of antibiotic resistance (AR), emerging trends, and novel resistance mechanism in clinical bacteria.

The perspective on the use of WGS are multiple and addressed to more fields as healthcare settings; for example, urinary catheters, human gut, skin for monitoring and investigating the human microbiome. At the same to investigate how human microbiome can develop and adapt influenced by environment such as food, antibiotics use, travel. WGS data can provide a clear link between isolates cases of food poisoning and environmental samples from premises associated with outbreaks [2]. Moreover, WGS data provide information on phylogenetic relationships – ancestral history.

Many other areas where WGS can have an impact in terms of public health setting from ongoing collaborations between researchers and clinicians. The use of WGS in routine laboratory work to decipher the resistome on bacterial strains or directly in the sample will be possible as costs fall, as the WGS becomes smaller and as databases improve. These innovative technologies are able to detect multi-drug resistant bacteria (MDRB) at routine practice to allow patient isolation, prevent nosocomial infections and the spread of bacteria, and adapt the therapeutic strategy as quickly as possible.

We can hope that in the future, using Next Generation Sequencing with lower cost and rapid phenotypic analysis can help to personalize therapies for individual treatment of infectious diseases, to improve food protection strategies, and to develop gene therapy. Moreover, we can use bacteria communities as biomarkers of the effectiveness of therapeutic intervention.

References

1. Galán J-C, González-Candelas F, Rolain J-M, Cantón R. Antibiotics as selectors and accelerators of diversity in the mechanisms of resistance: from the resistome to genetic plasticity in the β-lactamases world. *Front. Microbiol.* 4, 9 (2013).

2. King LA, Nogareda F, Weill F-X, *et al.* Outbreak of Shiga toxin-producing Escherichia coli O104:H4 associated with organic fenugreek sprouts, France, June 2011. *Clin. Infect. Dis.* [Internet]. 54(11), 1588–94 (2012). Available from: http://www.ncbi.nlm.nih.gov/pubmed/22460976.