

Université Paris-Est

HDR

Mémoire présenté en vue de l'obtention de l'habilitation à diriger des recherches

Spécialité : Sciences du vivant

par

Dr Nicolas Radomski

La génomique bactérienne
en appui à la santé publique

The bacterial genomics
in support of public health*

Soutenance : 13 mars 2020

Composition du jury

Dr Médigue Claudine	CNRS	Rapportrice#
Dr Desenclos Jean-Claude	Santé Publique France	Rapporteur#
Dr Franz Eelco	RIVM	Rapporteur#
Dr Moulin Laurent	Eau de Paris	Invité
Dr Chiapello Hélène	MaIAGE, INRAE	Examinatrice
Dr Glaser Philippe	Institut Pasteur	Examinateur
Dr Vergu Elisabeta	MaIAGE, INRAE	Examinatrice
Dr Brisabois Anne	ANSES (Directrice HDR)	Examinatrice
Dr Mistou Michel-Yves	INRAE (Directeur HDR)	Examinateur

* English synthesis in pages 78-85

Sélectionné(e) par l'Université Paris-Est

PRÉAMBULE

L'habilitation à diriger des recherches, a été créée en application de la loi du 26 janvier 1984 sur l'enseignement supérieur [1], qui a engendré plusieurs réglementations, en commençant par l'arrêté du 23 novembre 1988 [2], au moins suivit des deux circulaires des 05 janvier 1989 [3] et 27 octobre 1992 [4], puis des arrêtés des 13 février 1992 [5], 13 juillet 1995 [5], 25 avril 2002 [6], du 25 mai 2016 [7] et 1^{er} juillet 2016 [8].

L'arrêté du 23 novembre 1988 [2] a été modifié par les arrêtés des 13 février 1992 [9], 13 juillet 1995 [5] et 25 avril 2002 [6]. Cet arrêté du 23 novembre 1988 [2] est relatif à l'habilitation à diriger des recherches et stipulent que « *l'habilitation à diriger des recherches sanctionne la reconnaissance du haut niveau scientifique du candidat, du caractère original de sa démarche dans un domaine de la science, de son aptitude à maîtriser une stratégie de recherche dans un domaine scientifique ou technologique suffisamment large et de sa capacité à encadrer de jeunes chercheurs* ». D'après cet arrêté du 23 novembre 1988 [2], l'habilitation à diriger des recherches permet de candidater aux corps des professeurs des universités et directeurs de recherche.

L'arrêté du 25 mai 2016 [7], abroge certains articles de l'arrêté du 7 août 2006 relatif à la formation doctorale [10], a été modifié par l'arrêté du 1^{er} juillet 2016 [8] et fixe le cadre national de la formation et les modalités conduisant à la délivrance du diplôme national de doctorat. Cet arrêté du 25 mai 2016 [7] souligne que l'habilité à diriger des recherches peut participer à la désignation du directeur de l'école doctorale à laquelle il est inscrit, examiner en tant que rapporteur désigné par le chef d'établissement les travaux de thèse d'un(e) étudiant(e) en doctorat, ainsi que prendre la responsabilité d'un encadrement de thèse de doctorat en tant que directeur de thèse.

Le présent mémoire est proposé en vue de l'obtention de l'habilitation à diriger des recherches auprès de l'université Paris-Est qui demande au candidat une adéquation du champ disciplinaire à ceux couverts par les écoles doctorales, un niveau scientifique majeur mesuré par la qualité des productions, un intérêt pour l'université Paris-Est lorsque le candidat est extérieur à cet établissement, ainsi qu'un respect de la recevabilité des rapporteurs proposés, en vérifiant que ces derniers ne soient pas le directeur de thèse, le directeur d'habilitation du candidat ou toute personnalité ayant des liens d'intérêt ou ayant collaborée scientifiquement avec le candidat [11].

Ce mémoire commence par un chapitre présentant mon lieu d'accueil qui est un membre associé à l'université Paris-Est : l'agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail. Au regard de ce premier chapitre de mémoire et d'un second chapitre présentant mes travaux de recherche, un troisième chapitre est consacré à l'adéquation entre ma candidature à l'habilitation à diriger des recherches, mes travaux de recherche réalisés jusqu'à présent, et les projets que je souhaite diriger dans les prochaines années.

Dans un contexte de crises sanitaires, aux échelles nationale, européenne et internationale, les organismes de santé publique correspondants s'efforcent en synergie d'apporter des réponses aux populations dans de multiples domaines scientifiques. Personnellement engagé depuis plusieurs années dans des travaux de recherche en microbiologie dans les domaines de la santé humaine et chez l'animal, ainsi que la sécurité environnementale et alimentaire, je considère l'avènement des nouvelles technologies de séquençage des génomes, comme une opportunité pour la santé publique, d'améliorer ses systèmes de contrôle et prévention des risques sanitaires. Ceci est en particulier le cas dans le domaine de la microbiologie contemporaine qui est fortement modifiée par l'approche multisectorielle « *One Health* ».

REMERCIEMENTS

Je compte en premier lieu remercier mes directrice et directeur d'habilitation à diriger des recherches, Anne Brisabois et Michel-Yves Mistou, ainsi que mon directeur de laboratoire Laurent Laloux, pour m'avoir encouragé, guidé et conseillé concernant la présente candidature. Au regard de leurs apports scientifiques, merci aussi à mes anciennes et anciens responsables Michel-Yves Mistou, Marcel A. Behr, Régis Moilleron, Françoise Lucas, Laurent Moulin, Maria Laura Boschioli, Cécile Rousseau, Françoise Odelin, Graziella Bourdin, Pierre Malle, Jean-Pierre Vincent et David Corroler.

Merci aussi pour leurs collaborations à mes collègues scientifiques Arnaud Felten, Ludovic Mallet, Federica Palma, Pierre-Emmanuel Douarre, Maroua Sayeb, Déborah Merda, Emeline Cherchame, Alejandro Cabezas-Cruz, Thomas Texier, Pierre-Yves Letournel, Yannick Blanchard, Yannick Marie, Mathilde Bonis, Mahendra Mariadassou, Hélène Chiapello, Sabrina Cadel-Six, Patrick Fach, Sabine Herbin, Olivier Firmesse, Jacques Antoine Hennekinne, Jean-Charles Leblanc, Vincent Leclerc, David Albert, Renaud Lailier, Laurent Guillier, Thomas Brauge, Philippe Bessieres, Philippe Velge, Jacques Printems, Patrick Boiron, Veronica Rodriguez-Nava, Françoise Irlinger, Olivier Bezier, Catherine Lorgeoux, Lila Boudahmane, Nicole Buet, Uthaya Sivanantham, Emile Ducarme, Vincent Rocher, Sophie Haenn, Fanny Richard, Héberte Accrombessi, Claudine Wichlaz, Salak Gallah, Marie-Noël Gamard, Benjamin Richard, Gilles Varrault, Johnny Gasperi, Justin Tanner, Krista Williams, Myra Williams, Amy Pruden, Joseph O. Falkinham III, Frédéric Veyrier, Fiona McIntosh, Karl-Philippe Guérard, Damien Montamat-Sicotte, Louise Lefrançois, Pilar Domenech, Joyce Wang, Barbar Mindt, Michael B. Reed, Franck Biet, Christophe Sola, Emmanuelle Cambau, Benoit Cournoyer et Daniel Thévenot.

Les travaux de recherche présentés dans ce manuscrit n'auraient aussi pas pu se réaliser sans le concours de plusieurs étudiantes et étudiants en licence, master, doctorat et médecine, nommément Méryl Vila Nova, Marie-Noel Mansour, Kévin La, Pauline Barbet, Kévin Durimel, Lena Fritsch, Abdelrahim Abakabir-Mahamat, Yann Sévellec, Audrey Artignan, Robyn S. Lee, Trevor Tarakjian, Andrei Dan, Jeremy Dabor, Louis Kreitmann, Yacine Boudali et Laetitia Betelli. Merci aussi à d'autres étudiantes et étudiants en doctorat que j'ai pu suivre dans leurs travaux de recherche, Maëllys Kevin, Franck Cerutti et Damien Richard. Essentiels aux fonctionnements des laboratoires, je compte aussi remercier mes collègues des services administratifs, à savoir Catherine Delorme, Véronique Baum, Christelle Bendjedou, Christelle Gautheron, Nathalie Bongoua, Nathalie Dallet, Lynn Dery Capes, Catherine Charleux, Annick Piazza, Patricia Cambergs, Catherine Alcouffe et Marine Daniel. Ces travaux de recherche m'ayant conduit à travailler dans de nombreux laboratoires et émigrer aux États-Unis, puis au Canada, avant de revenir en France, je souhaite de plus assurément remercier mes parents Martine Radomski et Jean-Yves Radomski, ainsi que ma sœur Céline Radomski et mon frère Michaël Radomski, pour leurs encouragements indéfectibles.

À la fois profondément personnels et paradoxalement communautaires, les conceptions, financements, réalisations et publications de ces travaux de recherche sont à mettre à mon crédit et celui de l'ensemble des personnes y ayant contribué, y compris l'ensemble des équipes d'éditeurs et scientifiques qui ont permis de faire reconnaître et connaître ces résultats de recherche approuvés par l'évaluation par les pairs. C'est avec grand plaisir que je vous présente mes travaux de recherche co-signés de mes responsables, collègues et à ma grande satisfaction des étudiantes et étudiants y ayant contribué significativement. Pour finir, merci à l'ensemble des membres de mon jury : Claudine Médigue, Jean-Claude Desenclos, Eelco Franz, Laurent Moulin, Hélène Chiapello, Philippe Glaser, Elisabeta Vergu, Anne Brisabois et Michel-Yves Mistou.

« *La connaissance parle, mais la sagesse écoute* »

Jimi Hendrix

TABLES DES MATÈRES

PRÉAMBULE	4
REMERCIEMENTS	5
TABLES DES MATÈRES	7
ACRONYMES	13
LÉGENDES DES FIGURES	17
LÉGENDES DES TABLEAUX	18
LÉGENDES DES ANNEXES	19
ADRESSES WEB	20
INTRODUCTION	23
I. L'évolution de la microbiologie et des crises sanitaires.....	23
I.1. Point de vue des historiens	23
I.2. Point de vue des scientifiques.....	24
II. La numérisation des données de santé publique.....	26
II.1. Bouleversements technologiques	26
II.2. Coûts des collectes de données	26
II.3. Développement et implémentation de l'analyse de données	27
II.4. Modifications de l'organisation des laboratoires	27
III. Les enjeux du mémoire	30
III.1. Instructions légales et institutionnelles	30
III.2. Point de vue personnel.....	30
CHAPITRE I : LIEU D'ACCUEIL.....	33
I. L'agence d'accueil : Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail	33
I.1. Rayon d'activités	33
I.1.a. Tutelles	33
I.1.b. Missions.....	33
I.1.c. Champs de compétences	33
I.1.c.i. Surveillance et vigilance	34
I.1.c.ii. Expertise.....	34
I.1.c.iii. Recherche	34
I.2. Organisation.....	34
I.2.a. Fonctionnelle	34
I.2.a.i. Conseil d'administration	35
I.2.a.ii. Conseil scientifique.....	35
I.2.a.iii. Experts	35
I.2.a.iv. Laboratoires	35

I.2.b. Thématique.....	36
I.2.c. Hiérarchique	36
II. Laboratoire d'accueil : Laboratoire de sécurité des aliments	36
II.1. Rayon d'activités	37
II.1.a. Surveillance.....	37
II.1.b. Référence.....	37
II.1.c. Recherche	37
II.2. Organisation.....	38
II.2.a. Missions	38
II.2.b. Mandats.....	39
III. Équipe d'accueil : Mission génome analyse modélisation risque	39
III.1. Organisation.....	39
III.2. Objectifs	39
III.2.a. Typage	40
III.2.b. Attribution des sources	40
III.2.c. Recherche avancée en génomique.....	40
III.2.d. Diffusion de l'expertise	41
III.3. Supports.....	41
CHAPITRE II : BILAN DE CARRIÈRE	45
I. Contextes et financements des travaux.....	45
I.1. Santé humaine	45
I.2. Santé animale	47
I.3. Sécurité environnementale.....	47
I.4. Sécurité alimentaire	47
II. Encadrement et transversalité des travaux.....	48
II.1. Participation à des projets collaboratifs.....	48
II.1.a. Amélioration continue des développements	48
II.1.b. Intégration des développements à un réseau Linux	50
II.1.c. Transmissions des données et développements stables.....	52
II.2. Direction de projets propres	53
II.2.a. Travaux de recherche dirigés	53
II.2.b. Supports d'activités	60
II.2.b.i. Marchés publics et procédures afférentes.....	60
II.2.b.ii. Mise en place des moyens matériels.....	61
II.2.b.iii. Animation de la recherche	61
II.2.b.iv. Participation à la formation par la recherche	61

CHAPITRE III : PROGRAMME DE RECHERCHE	65
I. Adéquations entre environnement de travail et carrière personnelle	65
I.1. Implémentation de la génomique par les acteurs de la santé publique	65
I.1.a. Au sein des autorités de santé publique	65
I.1.b. Au sein du LSAL de l'ANSES	66
I.1.c. Au sein de la mission GAMeR	66
I.2. Amélioration des pratiques analytiques et intégration des domaines d'étude de la microbiologie	66
I.2.a. L'amélioration des pratiques analytiques en microbiologie	66
I.2.b. Intégration des domaines d'étude de la microbiologie	67
II. Projets de recherche	69
II.1. Projet de surveillance : Accessibilité des métadonnées des souches en collection au LSAL dans le cadre de l'intégration à la base de données génomique développée par la mission GAMeR (Projet AcSaMe)	69
II.1.a. Contexte scientifique	69
II.1.b. Travaux et financement envisagés	69
II.2. Projet de référence : Évaluation des déterminants de non-diminution de l'incidence de <i>Salmonella</i> Enteritidis dans le cadre de l'application de la génomique aux filières aviaires (Projet ADONIS)	70
II.2.a. Contexte scientifique	70
II.2.b. Travaux et financement envisagés	71
II.3. Projet de recherche : Évaluation à l'échelle génomique des sources hydriques à l'origine des TIAC à <i>Salmonella</i> dans l'union européenne (Projet FoWaBoSa) ..	71
II.3.a. Contexte scientifique	71
II.3.a.i. Pathogène alimentaire en Europe	71
II.3.a.ii. Pathogène alimentaire d'origine hydrique aux États-Unis	72
II.3.a.iii. Origine hydrique potentielle des TIAC en Europe ou en France	72
II.3.b. Travaux et financement envisagés	72
DISCUSSION	75
I. Investissements en capacités de stockage, ressources de calcul et compétences en bioinformatique	75
II. Accessibilité aux métadonnées épidémiques pour les développeurs et analystes en bioinformatique	75
III. Harmonisation de la génomique des pathogènes alimentaires pour l'efficacité de la santé publique	76
IV. Impact environnemental majeur de la génomique	76
CONCLUSION	76

SYNTHÈSE EN ANGLAIS.....	79
I. Host laboratory	79
I.1. French Agency for Food, Environmental and Occupational Health and Safety	79
I.2. Laboratory for food safety	79
I.3. Genome Analysis Modelling Risk	79
II. Career outcome	80
II.1. Human health	80
II.2. Animal health	80
II.3. Environmental safety.....	81
II.4. Food safety	81
II.4.a. Funding and human resources	81
II.4.b. Main genomic workflows	81
II.4.c. Computing and storage system	82
II.4.d. Collaborative research projects	82
II.4.e. GAMeR research projects	83
III. Future projects	83
III.1. Surveillance project: Accessibility to LSAL epidemiological metadata and integration in the genomic database developed by the GAMeR team (Project AcSaMe).....	83
III.1.a. Scientific context	83
III.1.b. Plan and funding	84
III.2. Reference project: Assessing determinants of the non-decreasing incidence of <i>Salmonella</i> Enteritidis (Project ADONIS).....	84
III.2.a. Scientific context	84
III.2.b. Plan and funding	84
III.3. Research project: Evaluation at the genomic scale of hydric sources causing <i>Salmonella</i> foodborne outbreaks in European Union (Project FoWaBoSa).....	85
III.3.a. Scientific context	85
III.3.b. Plan and funding	85
ANNEXES	87
RÉFÉRENCES	113
RÉSUMÉ	126

« Dans le domaine scientifique, trouver la bonne formulation d'un problème permet souvent de le résoudre »

Stephen Hawking

ACRONYMES

ABB : analyse biologique et biochimique
 ABIES : agriculture, alimentation, biologie, environnement, santé
 ACS : société chimique américaine
 AcSaMe : accessibility of sampling metadata
 ACTEOLab : application pour la centralisation et le transfert de données dédiées à l'épidémiosurveillance opérationnelle des laboratoires
 ADONIS : assessing determinants of the non-decreasing incidence of *Salmonella* Enteritidis
 AFEM : association française d'écologie microbienne
 AFLP : amplified fragment-length polymorphism
 AFSSA : agence française de sécurité sanitaire des aliments
 AFSSET : agence française de sécurité sanitaire de l'environnement et du travail
 AMR: antimicrobial resistance
 ANMV : agence nationale du médicament vétérinaire
 ANSES : agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail
 ANR : agence nationale de la recherche
 APHP : assistance publique hôpitaux de Paris
 API : interface de programmation d'applications
 ARCIR : actions de recherche concertées d'initiative régionale
 ARTwork : assembly of reads and typing workflow
 AVIESAN : alliance nationale pour les sciences de la vie et de la santé
 BfR : bundesinstitut für risikobewertung
 BIGSdb : bacterial isolate genome sequence database
 BsM : Boulogne-sur-Mer
 CDC : centers for disease control and prevention
 CDI : contrat à durée indéterminée
 CES: comités d'experts spécialisés
 CGE : center for genomic epidemiology
 cgMLST : coregenome MLST
 cgSNPs : coregenome SNPs
 COMPARE : collaborative management platform for detection and analyses of Re-emerging and foodborne outbreaks in Europe
 COMPASS : comprehensive and complete plasmid database
 CNRS : centre national de la recherche scientifique
 CPER : contrats de plan État-Région
 CRefIX : centre de référence, d'innovation, d'expertise et de transfert
 CTM : centre de typage moléculaire
 DAF : direction de l'administration et des finances
 DDBJ : DNA data bank of Japan
 DEBUG : débats autour de la bioinformatique et la génomique

DER : direction de l'évaluation des risques
DGAL : direction générale de l'alimentation
DIM : domaines d'intérêt majeur franciliens
DPI : déclaration publique d'intérêt
DTI : direction technique et informatique
DTU Food : national food institute, technical university of Denmark
DSP : direction de la stratégie et des programmes
ECDC : European centre for disease prevention and control
ED : écoles doctorale
EFSA : European food safety authority
EJP : European joint programme
EPA : établissement public à caractère administratif
ENCP : école des ponts ParisTech
ERPCB : équipe de recherche en physico-chimie et biotechnologie
ESA : épidémiosurveillance en santé animale
ESV : épidémiosurveillance en santé végétale
EU : European Union
FAO : organisation des nations unies pour l'alimentation et l'agriculture
fastGOES : fast gene ontology enrichment study
FBO : foodborne outbreak
FoWaBoSa : food and water-borne *Salmonella*
FP7 : seventh framework programme
FRQS : fonds de recherche en santé Québec
FranceAgriMer : établissement national des produits de l'agriculture et de la mer
FUI : fonds unique interministériel
GAMeR : génome analyse modélisation risque
GB : génie biologique
GMI : global microbial identifier
GOEA : gene ontology enrichment analysis
GSA : genome sequence archive
GTEvaluator : genotype evaluator
GTFinder : genotype finder
GWAS : genome wide association study
H2020 : horizon 2020
HACCP : hazard analyses critical control point
HDR : habilitation à diriger des recherches
IAA : industries agroalimentaires
IAB : industries alimentaires et biologiques
ICM : institut du cerveau et de la moelle épinière
ILIS : institut lillois d'ingénierie de la santé
InDels : small insertions/deletions
INNUENDO : a cross-sectorial platform for the integration of genomics in surveillance of food-borne pathogens
INRA : institut national de la recherche agronomique

INRAE : institut national de recherche pour l'agriculture, l'alimentation et l'environnement
 institut national de la recherche agronomique
 INSDC : international nucleotide sequence database collaboration
 IPL : institut Pasteur de Lille
 IRIDA : integrated rapid infectious disease analysis
 IUT : institut universitaire et technologique
 iVARcall2 : independent variant calling analysis
 LEESU : laboratoire eau environnement et systèmes urbains
 LHN : laboratoire d'hydrologie de Nancy
 LNR : laboratoires nationaux de référence
 LRUE : laboratoire de référence de l'union européenne
 LSAL : laboratoire de sécurité des aliments
 LSAN : laboratoire de santé animale
 MA : Maisons-Alfort
 MaIAGE : mathématiques et informatique appliquées du génome à l'environnement
 MAGB : microbiologie appliquée et génie biologique
 MAPA : marché à procédure adaptée
 microbial-GWAS : microbial genome wide association study
 MIRU : mycobacterial interspersed repetitive unit
 MLST : multilocus sequence typing
 MNT : mycobactéries non-tuberculeuses
 MycoClub : société française de microbiologie
 NCBI : national center for biotechnology information
 NGS : next generation sequencing
 NIFDS : national institute of food and drug safety evaluation
 NIPH: Norwegian institute of public health
 NTM : non-tuberculous mycobacteria
 ORCID : open researcher and contributor identifier
 OIE : organisation mondiale de la santé animale
 OMS : organisation mondiale pour la santé
 ONU : organisation des nations unies
 OPUR : observatoire des polluants urbains
 Pairwise-FBO : Pairwise foodborne outbreaks
 PCR : polymerase chain reaction
 PFGE : pulsed-field gel electrophoresis
 PHAC : agence de la santé publique du Canada
 PhyloFixedVar : Phylogeny fixed variants
 PIREN-Seine : programme interdisciplinaire de recherche sur l'eau et l'environnement du bassin de la Seine
 PIWET : państwowy instytut weterynaryjny
 RaDAR: risk and disease burden of antimicrobial resistance
 RFLP : restriction fragment length polymorphism
 RI MUHC : institut de recherche du centre de santé de l'université de McGill
 RIVM : rijksinstituut voor volksgezondheid en milieu

RRSSSN : régie régionale de la santé et des services sociaux du Nunavik
SBCL : *Staphylococcus, Bacillus, Clostridium*
SCA : plateforme de surveillance sanitaire de la chaîne alimentaire
SEL : *Salmonella* et *Listeria*
SERMHA : service de microbiologie et hygiène alimentaire
SFBI : société française de bioinformatique
SFM : société française de microbiologie
SIAPP : syndicat interdépartemental pour l'assainissement de l'agglomération parisienne
SMRE : sciences de la matière, du rayonnement et de l'environnement
SNPs : single nucleotide polymorphisms
SPI : *Salmonella* pathogenic islands
SSI : statens serum institut
TIAC : toxi-infections alimentaires collectives
UE : union européenne
UPE : université Paris-Est
UPEC : université Paris-Est de Créteil
UZB : unité de zoonoses bactériennes
VENoMYC : veterinary network of laboratories researching into improved diagnosis and epidemiology of mycobacterial diseases
VNTR : variable number tandem repeat
VPN : virtual private network
wgMLST : whole genome MLST
WGS : whole genome sequencing
WP : working package

LÉGENDES DES FIGURES

Figure 1: Points de vue des historiens et scientifiques sur l'histoire de la microbiologie et les crises sanitaires majeures.....	25
Figure 2 : Éléments modifiés dans l'organisation des laboratoires de « Public Health England » (PHE) avant et après l'implémentation du « whole genome sequencing » (WGS) des pathogènes alimentaires majeurs (<i>Salmonella</i> , <i>Escherichia coli</i> , <i>Shigella</i> , <i>Listeria</i> et <i>Campylobacter</i>).....	29
Figure 3 : Logos français (à gauche) et anglais (à droite) de l'agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (ANSES)	34
Figure 4 : Laboratoires (cercles) et antennes de laboratoires (carrés) de l'agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (ANSES) à Angers, Boulogne-sur-Mer, Clermont-Ferrand, Dozulé, Fougères, Maisons-Alfort, Malzéville, Montpellier, Nancy, Niort, Ploufragan, Plouzané, Rennes, Saint-Pierre de la Réunion et Sophia-Antipolis	35
Figure 5 : Domaines d'activité du laboratoire de sécurité des aliments (LSAL) sur les sites de Maisons-Alfort et Boulogne-sur-Mer	38
Figure 6 : Logos français (à gauche) et anglais (à droite) de la mission génome analyse modélisation risque (GAMeR).....	39
Figure 7 : Pouvoir discriminant de méthodes de caractérisation pré- et post-next generation sequencing (NGS) : sérotypage, ribotyping, « amplified fragment-length polymorphism » (AFLP), « restriction fragment length polymorphism » (RFLP), « variable number tandem repeat » (VNTR), « multilocus sequence typing » (MLST), « pulsed-field gel electrophoresis » (PFGE), « coregenome MLST » (cgMLST), « whole genome MLST » (wgMLST) et « coregenome SNPs » (cgSNPs)	40
Figure 8 : Outil « independent variant calling analysis » (iVARcall2) développé par la mission génome analyse modélisation risque (GAMeR) au sein du laboratoire de sécurité des aliments (LSAL) pour identifier les variants de coregénomés bactériens par « mapping » et « variant calling » [96, 97, 99, 101, 102, 104, 118, 120, 128]	49
Figure 9 : Outil « assembly of reads and typing workflow » (ARTwork) développé par la mission génome analyse modélisation risque (GAMeR) au sein du laboratoire de sécurité des aliments (LSAL) pour assembler <i>de novo</i> des génomes bactériens [96–99, 102, 104, 128]	49
Figure 10 : Réseau Windows-Linux développé par la mission génome analyse modélisation risque (GAMeR) au sein du laboratoire de sécurité des aliments (LSAL) en collaboration avec la direction technique et informatique (DTI), la plateforme génomique de Ploufragan et l'institut du cerveau et de la moelle épinière (ICM).....	51
Figure 11 : Évolution du nombre de génomes intégrés à la base de données de la mission génome analyse modélisation risque (GAMeR) provenant des genres bactériens d'archives internationales ou isolés et collectés au laboratoire de sécurité des aliments (LSAL) : <i>Clostridium</i> (n = 216), <i>Bacillus</i> (n = 283), <i>Staphylococcus</i> (n = 316), <i>Listeria</i> (n = 2 605) et <i>Salmonella</i> (n = 2 919) en date du 10 juillet 2019 (n = 6 339).....	53
Figure 12 : Inférences phylogénomiques des SNPs du coregénome de souches de <i>Salmonella enterica</i> spp. <i>enterica</i> (n = 192 génomes) impliquées dans des épidémies des serovars <i>S. Typhimurium</i> (A) and <i>S. 1,4,[5],12:i:-</i> (B) et résultats des tests non-paramétriques Wilcoxon rank sum (WS), Kolmogorov-Smirnov (KS) and Kruskal-Wallis (KW) [92]	55
Figure 13 : Inférence phylogénomique des SNPs du coregénome de souches de <i>Salmonella enterica</i> spp. <i>enterica</i> (n = 73 génomes) des serovars Dublin, Enteritidis, Pullorum et Gallinarum, respectivement des modèles d'adaptation aux bovin, à de multiples hôtes et des hôtes aviaires, visant à identifier les variants fixés aux nœuds des divergences (c.à.d. l'outil PhyloFixedVar) [101].....	56

Figure 14 : Inférence phylogénomique des SNPs du coregénomique de souches de <i>Salmonella enterica</i> spp. <i>enterica</i> (n = 440 génomes) de serovars considérés comme adaptés à des sources animales mono- et multi-hôtes et mutations associées par « genome wide association study » (GWAS) (c.à.d. l'outil microbial-GWAS) [102].....	57
Figure 15 : Mutations du génome accessoire et du coregénomique associées par « genome wide association study » (GWAS) (c.à.d. l'outil microbial-GWAS) à des sources animales mono- et multi-hôtes de sérovars de <i>Salmonella enterica</i> spp. <i>enterica</i> (n = 440 génomes) et voies métaboliques correspondantes identifiées par « gene ontology enrichment analysis » (GOEA) (c.à.d. l'outil fastGOES) [102].....	58
Figure 16 : Distribution des tailles de plasmides (log ₁₀ de paires de bases) en fonction des Phyla bactériens de la base de plasmides (n = 12 084) « comprehensive and complete plasmid database » (COMPASS) [103]	59
Figure 17 : Membres de la mission génome analyse modélisation risque (GAMeR) en juillet 2018 (de gauche à droite et de haut en bas : Kévin La, Julien Pichon, Méryl Vila Nova, Arnaud Felten, Kévin Durimel, Émilie Pasteur, Pauline Barbet, Pierre-Emmanuel Douarre, Ludovic Mallet, Michel-Yves Mistou, Nicolas Radomski).....	62
Figure 18 : Articles publiés pendant mon parcours professionnel international orienté vers l'amélioration des outils analytiques en bactériologie, biologie moléculaire et génomique bactérienne dans les domaines de la microbiologie humaine, alimentaire, environnementale et alimentaire (le signe * signifie que le manuscrit est en cours d'évaluation par les pairs).....	68
Figure 19 : Interface de programmation (API) de l'EFSA fournissant des métadonnées et résultats analytiques aux développeurs en bioinformatique <i>versus</i> base de données génomiques gérée par la mission génome analyse modélisation risque (GAMeR) qui n'intègre pas actuellement de métadonnées épidémiques pour les analystes en bioinformatique.....	70
Figure 20 : Évolution dans l'union européenne (2008 et 2016) de la prévalence dans la filière des poules pondeuses (%) et des notification de cas humains (pour 100 000) de <i>Salmonella enterica</i> spp. <i>enterica</i> sérovar Enteritidis [89].....	71
Figure 21 : Situations contrastées entre le pathogène alimentaire <i>Salmonella</i> en Europe (2015-2017) et le pathogène alimentaire <i>Salmonella</i> d'origine hydrique aux Etats-Unis (2005-2013).....	73

LÉGENDES DES TABLEAUX

Tableau 1 : Initiatives nationale et internationale de plateformes analytiques en génomique des microorganismes pathogènes et autres microorganismes d'intérêt isolés chez l'Homme, d'animaux, de l'alimentation humaine et animale, ainsi que d'environnement de production alimentaire.....	42
Tableau 2 : Domaines de la microbiologie (santé humaine et animale, sécurité environnementale et alimentaire), cursus, lieux d'accueil (établissement, laboratoire, équipe), sujets de recherche et sources de financement	46
Tableau 3 : Outils analytiques développés par la mission génome analyse modélisation risque (GAMeR) entre 2014 et 2019.....	54
Tableau 4 : Prestations du marché à procédure adaptée (MAPA) portant sur le séquençage haut débit (Librairies Nextera XT et séquenceur Illumina 2 x 150 bases Nextseq500) de génomes bactériennes (taille de 2-7 Mb et couverture minimale théorique de 30X), conclu en 2018 (n°XBDC000257-18FCS024) entre l'agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (ANSES) et l'institut du cerveau et de la moelle épinière (ICM), valable pour 2 ans, et renouvelable deux fois pour une période totale de 6 ans.....	60

LÉGENDES DES ANNEXES

Annexe 1 : <i>Curriculum Vitae</i> détaillé.....	87
Annexe 2 : Liste des établissements pouvant établir des relations conventionnelles avec l'agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (ANSES).....	97
Annexe 3 : Organigramme hiérarchique de l'agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (ANSES) au 1 ^{er} avril 2019	98
Annexe 4 : Plaquette de présentation des trois plateformes d'épidémiologie et de leurs acteurs sous l'égide du ministère des solidarités et de la santé et du ministère de l'agriculture et de l'alimentation.....	99
Annexe 5 : Mandats de laboratoire national de référence (LNR : cercle blanc) et laboratoire de référence de l'union européenne (LRUE : cercle noir) des sites Maisons-Alfort (MA) et Boulogne-sur-Mer (BsM) du laboratoire de sécurité des aliments (LSAL) organisé en départements, unités et missions transversales (T)	101
Annexe 6 : Note d'organisation de la mission génome analyse modélisation risque (GAMeR) - LSA-NO-0023 - LSA-MM-0001 - Manuel Qualité A.....	102
Annexe 7 : Outil « genotype evaluator » (GTEvaluator) développé par la mission génome analyse modélisation risque (GAMeR) au sein du laboratoire de sécurité des aliments (LSAL) pour estimer la justesse (spécificité et sensibilité) de cibles moléculaires à l'échelle génomique [95].....	104
Annexe 8 : Outil « genotype finder » (GTFinder) développé par la mission génome analyse modélisation risque (GAMeR) au sein du laboratoire de sécurité des aliments (LSAL) pour identifier des kmers caractéristiques (spécifiques et sensibles) de sous-ensembles de génomes [91].....	105
Annexe 9 : Outil (A) « Pairwise foodborne outbreaks » (Pairwise-FBO) et principe statistique (B) développés par la mission génome analyse modélisation risque (GAMeR) au sein du laboratoire de sécurité des aliments (LSAL) pour assigner des génomes à des toxi-infections alimentaires collectives (TIAC) sur la base de tests non-paramétriques (Kolmogorov-Smirnov, Wilcoxon-Mann-Whitney, Kruskal-Wallis) comparant les différences par paires de SNPs (avec ou sans recombinaisons : SNP-1 et SNP-2), gènes, kmers et allèles (cgMLST et wgMLST) [92].....	106
Annexe 10 : Outil « Phylogeny fixed variants » (PhyloFixedVar) développé par la mission génome analyse modélisation risque (GAMeR) au sein du laboratoire de sécurité des aliments (LSAL) afin d'identifier les variants du coregénomique (SNPs and InDels), homoplastiques ou non-homoplastiques, fixés aux nœuds de l'inférence phylogénomique d'une population clonale ou panmictique, puis les voies métaboliques principalement enrichies par « gene ontology enrichment analysis » (GOEA) [101]	107
Annexe 11 : Outil « microbial genome wide association study » (microbial-GWAS) développé par la mission génome analyse modélisation risque (GAMeR) au sein du laboratoire de sécurité des aliments (LSAL) afin d'identifier les mutations (gènes accessoires et variants du coregénomique provenant d'événements de recombinaisons homologues ou non) associées à un trait phénotypique binaire dans une population clonale ou panmictique par « genome wide association study » (GWAS) [102]	108
Annexe 12 : Outil « fast gene ontology enrichment study » (fastGOES) développé par la mission génome analyse modélisation risque (GAMeR) au sein du laboratoire de sécurité des aliments (LSAL) afin d'identifier les voies métaboliques principalement enrichies par « gene ontology enrichment analysis » (GOEA) [102].....	109
Annexe 13 : Procédures de soumission d'échantillons à la mission génome analyse modélisation risque (GAMeR) mises en place entre 2014 et 2019 pour appliquer le séquençage haut débit, l'assemblage <i>de novo</i> et l'identification de variants de génomes bactériens	110

ADRESSES WEB

- Adresse web 1 : <https://www.anses.fr/fr>
- Adresse web 2 : http://www.nicolas-radomski.net/pdf/CV_Nicolas_Radomski_FR.pdf
- Adresse web 3 : <http://www.nicolas-radomski.net/pdf/publications-HDR-Nicolas-Radomski.zip>
- Adresse web 4 : <https://orcid.org/0000-0002-7480-4197>
- Adresse web 5 : <http://sas-vp-lsdb1:3000/>
- Adresse web 6 : <\\sas-vp-lsgw1>
- Adresse web 7 : <https://openapi-portal.efsa.europa.eu/>
- Adresse web 8 : <https://www.czbiohub.org/projects/infectious-disease/>
- Adresse web 9 : https://www.efsa.europa.eu/en/interactive_pages/Pesticides_report_2017
- Adresse web 10 : https://www.efsa.europa.eu/en/interactive_pages/AMR_Report_2015
- Adresse web 11 : <https://www.cdc.gov/ncbddd/hemophilia/communitycounts/data-viz.html>
- Adresse web 12 : <https://geodes.santepubliquefrance.fr>
- Adresse web 13 : http://www.nicolas-radomski.net/pdf/CV_Nicolas_Radomski_EN.pdf

*« Lorsqu'un homme s'installe avec un travail dans un coin,
il abandonne autant de vie qu'il acquiert de connaissance »*

Williams Butler Yeats

INTRODUCTION

Dans le cadre d'une microbiologie contemporaine tendant à intégrer de plus en plus de nouvelles technologies pour résoudre des crises sanitaires liées à des bouleversements d'écosystèmes (intro. § I), et d'une santé publique de plus en plus numérisée (intro. § II), le présent mémoire vise à candidater à l'habilitation à diriger des recherches (HDR) en tenant compte des instructions légales et institutionnelles (intro. § III).

I. L'évolution de la microbiologie et des crises sanitaires

D'un point de vue historique, la microbiologie pourrait être segmentée en quatre phases : la découverte de la microscopie au XVII^{ème} siècle, les travaux en bactériologie et parasitologie au XIX^{ème} siècle, l'essor de la biologie moléculaire au XX^{ème} siècle, et la révolution des disciplines omiques au XXI^{ème} siècle (intro. § I.1) (Figure 1). D'un point de vue scientifique, les maladies infectieuses pourraient avoir évolué en cinq crises sanitaires humaines majeures induites par des modifications profondes d'écosystèmes (intro. § I.2) (Figure 1).

I.1. Point de vue des historiens

Le père de la microbiologie au XVII^{ème} siècle, le savant néerlandais Antonie van Leeuwenhoek, en raison de sa propre conception des microscopes, a été le premier à observer des microorganismes, à savoir en décrivant ce que nous nommons aujourd'hui des protozoaires en 1674 [12], des algues microscopiques en 1676 [13] et des spermatozoïdes en 1677 qu'il nomme animacules [14] ou « *animacula* » [12]. Ne rejetant pas totalement la théorie de la génération spontanée, Antonie van Leeuwenhoek conclut que ces animacules ou leurs semences étaient préexistants dans l'eau de pluie [15]. Ce n'est que 200 ans plus tard que la microbiologie a de nouveau été révolutionnée par les travaux de Louis Pasteur en France sur les cultures pures (lactique et butyrique) [16], la réfutation de la théorie de la génération spontanée [16], la dissymétrie moléculaire (l'acide tartrique) [17], les maladies des vins français (*Acetobacter aceti*) [18] et la maladie des vers à soie dans les Cévennes (*Nosema bombycis*) [19]. Louis Pasteur a aussi travaillé sur la pasteurisation alimentaire [18], l'asepsie en milieu médical [20], ainsi que les vaccinations contre le choléra des poules (*Pasteurella avicida*) [21], la maladie du charbon chez les moutons (*Bacillus anthracis*) [22], le rouget des porcs (*Erysipelothrix rhusiopathiae*) [23] et la rage (*Lyssavirus*) [24]. Durant cette même seconde moitié du XIX^{ème} siècle, le médecin allemand fondateur de la bactériologie, le professeur Robert Koch, découvrait quant à lui des agents infectieux humains, notamment les bacilles du charbon (*Bacillus anthracis*) [25], de la tuberculose (*Mycobacterium tuberculosis*) [26] et du choléra (*Vibrio cholerae*) [27]. Robert Koch étudiait aussi la malaria (*Plasmodium falciparum*), la peste (*Yersinia pestis*), la theilériose bovine (*Theileria parva*), la babesiose (*Babesia divergens*), la tuberculose bovine (*Mycobacterium bovis*) [28] et la maladie à tique [29]. Après avoir émis son postulat de Koch sur les liens de cause à effet entre les microorganismes et les maladies [30], Robert Koch a aussi développé la tuberculine contre le bacille de la tuberculose [31]. La fin du XIX^{ème} et le début du XX^{ème} siècle ont aussi été marqués par le médecin biologiste et pharmacologue britannique, Alexander Fleming, pour ses découvertes du lysozyme en 1922 [32] et de la pénicilline en 1928 [33] qui ont révolutionné la médecine. Au début du XX^{ème} siècle, les travaux du microbiologiste italo-américain prix Nobel en 1969, Salvador Luria, sur l'apparition aléatoire de bactéries mutantes conférant une résistance à un virus, ont contribué à la naissance de la biologie moléculaire [34]. Issue de l'union entre la génétique et la biochimie, la biologie moléculaire est partie prenante de la microbiologie au XX^{ème} siècle, notamment avec la description de la structure en double hélice de l'ADN en 1953 par les prix Nobel de

physiologie ou médecine James Watson et Francis Crick [35] et l'invention de la technique « polymérase chain reaction » (PCR) en 1983 par les prix Nobel de chimie Kary Mullis et Michael Smith [36, 37]. La microbiologie du XXI^{ème} siècle est quant à elle modifiée actuellement en profondeur par l'avènement des disciplines omiques (intro. § II).

1.2. Point de vue des scientifiques

En accord avec François Renaud, médaille d'argent 2010 du centre national de la recherche scientifique (CNRS), l'Homme a déjà connu plusieurs crises sanitaires en raison de modifications de ses écosystèmes, y compris la confrontation à de nouveaux microorganismes pathogènes. Une première crise sanitaire a eu lieu au Néolithique il y a 12 000 ans, lorsque la migration à travers l'Afrique et l'Europe accompagnée de la domestication et de l'élevage, ont provoqué des transmissions d'homme à homme de la tuberculose (*Mycobacterium tuberculosis*) [38], ainsi que des transmissions à l'homme de pathogènes du bétail impliqués dans la diphtérie (*Corynebacterium diphtheriae*) [39], la variole (*Variola virus*) [40], la coqueluche (*Bordetella pertussis*) [39] et la rougeole (*Morbillivirus*) [41]. Par la suite, une seconde crise sanitaire a frappé l'humanité entre l'an 500 av. J.-C et le moyen âge, où le commerce avec l'Asie en développement, ainsi que les guerres et sièges de villes antiques ou médiévales, ont induit les grandes épidémies de peste (*Yersinia pestis*) en Europe, notamment la peste d'Athènes où périra Périclès en l'an 430 av. J.-C [42], la peste des Carthaginois assiégeant Syracuse en l'an 395 av. J.-C [43], la peste de Justinien du I^{er} siècle [44] et les pestes noire des VI et XV^{ème} siècles [45]. De multiples épisodes de lèpre (*Mycobacterium leprae*) ont aussi eu lieu pendant cette même époque s'étendant sur quinze siècles apr. J.-C et furent rapportés dans les chapitres XIII et XIV du Lévitique [46], ainsi que par des publications contemporaines traitant d'échantillons médiévaux [47]. La troisième crise sanitaire de l'humanité correspond à l'arrivée en Amérique du Sud à la fin du XV^{ème} siècle des Européens qui ont transmis de multiples pathologies infectieuses, comme par exemple la syphilis (*Treponema pallidum*), aux populations indigènes qui ont été décimées en raison du fait qu'elles étaient dépourvues de réponses immunitaires adéquates [48]. D'après l'encyclopédie Canadienne, la colonisation des européens à la découverte des Amériques a aussi apporté la grippe, la rougeole et la variole qui ont fait des ravages dans les populations autochtones [49]. La quatrième crise sanitaire humaine fut induite par l'industrialisation, le développement des villes et la précarité qui ont favorisé la réémergence de pathologies infectieuses entre les XIX^{ème} et XX^{ème} siècles, comme notamment le choléra (*Vibrio cholerae*) [50] et la tuberculose (*Mycobacterium tuberculosis*) [51]. La cinquième crise sanitaire humaine observée en ce début du XXI^{ème} siècle, est ponctuée par des épidémies de chikungunya imputées à l'importation de volailles [52], d'Ébola potentiellement associées à la déforestation [53], de Zika provoquant des macrocéphalies chez les nouveaux nés [54], de syndrome respiratoire aigu sévère [55] et de grippe aviaire [56]. Sans explications bien définies concernant les origines de l'ensemble de ces événements, certains auteurs émettent d'ailleurs une possible association entre la trajectoire du cyclone El Niño et les épidémies récentes de chikungunya, hantavirus, fièvre de la vallée du Rift, cholera, peste et Zika [57]. Dans un contexte d'augmentation exponentielle mondiale du prix de l'alimentation depuis les années 2000 [58], certains des événements épidémiques de cette cinquième crise sanitaire humaine du XXI^{ème} siècle sont probablement induits par les élevages intensifs [59] et échanges alimentaires massifs [60] devant répondre à la nécessité de nourrir 9,7 milliards d'humains en 2050 selon l'organisation des nations unies (ONU). Sur la base des découvertes historiques majeures en microbiologie évoquées précédemment (intro. § I.1) et afin de résoudre les multiples formes de crises sanitaires humaines contemporaines (intro. § I.2), la gestion des données en santé publique s'oriente aujourd'hui mondialement vers une numérisation globale (intro. § II).

L'histoire de la microbiologie

Les crises sanitaires majeures

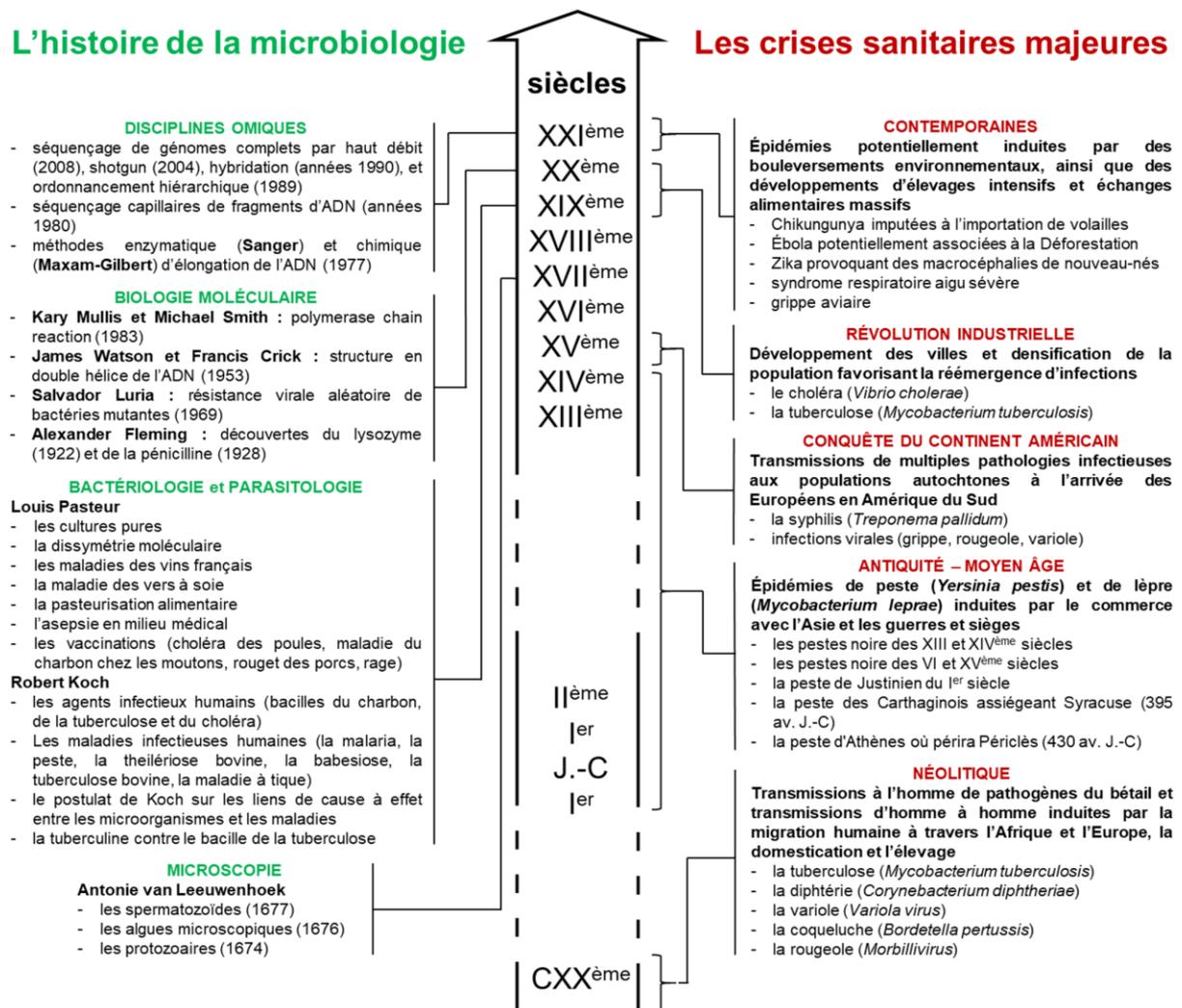


Figure 1: Points de vue des historiens et scientifiques sur l'histoire de la microbiologie et les crises sanitaires majeures

II. La numérisation des données de santé publique

En raison de la diminution des coûts de séquençage de l'ADN (intro. § II.1) et de l'augmentation des données devant être stockées dans les archives internationales (intro. § II.2), les dépenses engendrées par les disciplines omiques (génomique, transcriptomique, métagénomique, protéomique, lipidomique, épigénétique ...) sont aujourd'hui davantage dédiées au stockage et l'analyse de données (intro. § II.3), ainsi qu'à la réorganisation des laboratoires (intro. § II.4).

II.1. Bouversements technologiques

Depuis l'avènement du séquençage de petits fragments d'ADN par les méthodes d'élongation ou de dégradation de l'ADN en 1977, respectivement et communément connue sous les noms de Sanger [61] et Maxam-Gilbert [62], des équipements de séquençage capillaire sont apparus au début des années 1980 [63] et sont toujours d'actualité malgré leurs faibles débits [64]. Depuis ces années 80 et jusqu'à présent, le séquençage de génomes entiers, nommé aujourd'hui « whole genome sequencing » (WGS), a bouleversé ce domaine scientifique. Effectivement, sont apparues des méthodes dites par ordonnancement hiérarchique en 1989 [65], par hybridation dans les années 90 [66], par coup de fusil aussi connue sous le nom de « Shotgun » en 2004 [67] et par haut débit communément appelée « next generation sequencing » (NGS) ou « high-throughput sequencing » (HTS) depuis 2008 [68]. Alors que le séquençage du génome humain a pris 15 ans pour un coût de 100 million \$ US par la méthode Sanger [69], ce dernier a été séquencé en 2018 par NGS en deux mois pour un coût approximativement divisé par 100 [68]. La démocratisation de ces méthodes NGS a amené les constructeurs à proposer aujourd'hui différents modèles de séquenceurs en fonction du domaine d'expertise et des rendements analytiques souhaités : Illumina (MiniSeq, MiSeq, NextSeq, HiSeq, NovaSeq), IonTorrent (PGM, S5, Proton), Pacific BioSciences (PacBio RSII, Sequel) et Oxford Nanopore (MinION, GridION, PromethION) [70]. L'exemple le plus frappant de l'essor du NGS est la Chine qui prévoit de séquencer 100 millions de génomes de toutes espèces d'ici 2030 avec un budget de 9,2 milliards \$ US (60 milliards ¥ CN) sur 15 ans [71].

II.2. Coûts des collectes de données

Cette diminution drastique du coût du NGS a été accompagnée par une augmentation massive et exponentielle des données de séquençage dans les archives internationales qui selon la « international nucleotide sequence database collaboration » (INSDC) sont passées de 1,432 trillions de bases en août 2015 à 2,650 trillions de bases en août 2017 [72]. Cette augmentation de données de séquençage correspondait pour cette période de deux ans à une augmentation d'espace de stockage passant de 1,5 à 3,2 petabytes dédiées à la conservation d'une copie unique de données de séquençage non assemblées [72]. Depuis maintenant 30 ans et sous la gouvernance de la INSDC, ces afflux de données de séquençage ont donc conduit les bases de données de séquençage d'ADN américaine, européenne et japonaise à harmoniser et synchroniser leurs systèmes de stockage de données [72]. Les bases de données correspondantes « national center for biotechnology information » (NCBI), « european nucleotide archive » (ENA) et « DNA data bank of Japan » (DDBJ), sont principalement dédiées à la collecte et mise à disposition de données de séquençage et d'assemblage [72]. De plus, ces bases de données internationales proposent aussi quelques outils analytiques, principalement dédiés à la recherche de motifs nucléiques [72]. Plus précisément, l'ENA est hébergée par le « european molecular biology laboratory » (EMBL) et la « european bioinformatics institute » (EBI). En plus du coût matériel du stockage de l'information, un récent test de l'EMBL-EBI souligne que la nécessité d'organiser des ontologies et systèmes d'interrogation des métadonnées liées aux données de séquençage, présente aussi un coût important induit par les ressources de calcul nécessaires aux requêtes, mises à jours, rééquipement et charge salariale [73]. À titre d'exemple, ce test

d'annotation de 1 939 enregistrements de séquences avec les ressources de l'EMBL-EBI a nécessité 794 heures de charge salariale dédiée à l'amélioration de l'information contextuelle [73]. Ceci représenterait 37 personnes à temps plein pendant une semaine pour annoter l'ensemble des enregistrements de l'ENA en 2016, à raison de 100 échantillons par semaines [73]. En 2015, les coûts pour les utilisateurs des services libres de stockage et requêtes de l'EMBL-EBI étaient estimés annuellement à 57 millions \$US pour l'investissement, 329 millions \$US pour l'accessibilité et 2,8 milliards \$US pour l'usage (respectivement 47 millions £, 270 millions £ et 2,3 milliards £) [74]. En raison du développement contemporain massif du séquençage en Chine, la INSDC vient d'ailleurs de proposer un nouveau système d'archive de séquence nommé « genome sequence archive » (GSA) [75]. Bien que moins avancée que les systèmes d'archive des données produites par le NGS, il est à noter que la communauté scientifique de l'imagerie médicale commence à constituer ses instances publiques internationales d'archivage de données [76]. L'archivage de données est aussi considérée dans le domaine de l'imagerie médicale comme le goulot d'étranglement de la recherche biomédicale moderne [76].

II.3. Développement et implémentation de l'analyse de données

Le coût du séquençage ayant diminué drastiquement (intro. § II.1) et le coût du stockage restant à être résolu (intro. § II.2), les efforts actuels portent principalement sur le développement et l'implémentation de l'analyse de données dans différents domaines de la santé publique (intro. § II.3). Visant à appliquer la génomique pour 60 millions de patients entre 2020 et 2025 [77], l'intégration de la génomique aux systèmes hospitaliers a déjà commencé dans 14 pays sur la base d'investissements à hauteur de 4 milliards \$ US, en particulier pour des patients à maladies rares et atteints de cancer (Royaume-Uni, France, Australie, Arabie Saoudite et Turquie), des études de population (Etats Unis, Estonie, Danemark, Japon et Qatar) ou uniquement pour l'infrastructure informatique (Suisse, Hollande, Brésil et Finlande) [71]. Dans le domaine de l'étude de la biodiversité (sécurité environnementale), le projet « Tara oceans » est à souligner quant à son objectif ambitieux de cartographier les faune et flore microbiennes des océans par génomique et métagénomique [78]. En plus des connaissances apportées sur les microbiomes benthique et pélagique, les efforts du projet « Tara oceans » visent aussi à apporter des solutions analytiques au milieu médical, à l'élevage terrestre et à l'alimentation en générale, tant pour ce qui est des diagnostics que des contrôles dont les besoins sont en constante augmentation [79]. Un récent rapport de 2019 compare des initiatives contemporaines de développements d'outils analytiques sur la base des synthèses des autorités correspondantes : les agences « European centre for disease prevention and control » (ECDC) et « European food safety authority » (EFSA) [80]. Ce rapport focalise sur l'analyse génomique de données recueillies sur les microorganismes pathogènes et autres microorganismes d'intérêt isolés de l'Homme, d'animaux, de l'alimentation humaine et animale, ainsi que d'environnement de production alimentaire [80]. Dans ce rapport, l'ECDC et l'EFSA, responsables de la prévention de maladies infectieuses respectivement humaine et d'origine alimentaire, soulignent la nécessité de ressources de calcul pouvant être délivrées par différents opérateurs de service (Amazon, Microsoft, Google) [80]. L'ECDC et l'EFSA comparent aussi dans ce rapport les initiatives analytiques en génomique en terme de collection de données, assemblage, classification, nomenclature, contrôle et infrastructure, sans pour autant donner de recommandations particulières concernant leurs utilisations [80].

II.4. Modifications de l'organisation des laboratoires

Aujourd'hui, la capacité des laboratoires à séquencer des génomes complets à moindre coût (intro. § II.1), les conduit à focaliser leurs efforts sur les performances de stockage de données génomiques (intro. § II.2), ainsi que le développement et l'implémentation de l'analyse de données (intro. § II.3), ce qui engendre une profonde modification de leurs organisations (intro. § II.4). Un exemple intéressant à suivre de ces

modifications organisationnelles est celui mis en place par l'instance de santé publique anglaise, « Public Health England » (PHE), qui a publié comment ses laboratoires se sont transformés pour implémenter complètement la génomique des pathogènes alimentaires majeurs en 6 ans de 2012 à 2017 [81]. Effectivement, avec des objectifs d'identification de pathogènes, d'estimation de proximités génétiques, d'identification de résistances aux antibiotiques et caractères de virulence, ainsi que d'investigations d'épidémies, PHE a commencé à construire un système en 2012 [81]. PHE est passée de 7 640 à 15 750 génomes traités annuellement entre 2014 et 2017, pour atteindre un total en 2017 de 52 590 génomes de *Salmonella*, *Escherichia coli*, *Shigella*, *Listeria* et *Campylobacter* [81]. PHE est passé d'une organisation dépendante des pathogènes analysés accompagnée des multiples méthodes de biologie moléculaire correspondantes, à une organisation centralisée et divisée en une partie *in vitro* et une partie *in silico* traitant respectivement le séquençage et l'analyse de données (« wetlab » et « drylab ») [81]. PHE a pu passer de 30 jours à 5 jours de délais analytique (PCR + serotypage + caractérisation de phage + génotypage *versus* séquençage + analyse de données) avec 1 700 TB de stockage et une puissance de calcul de 656 cores distribués sur 34 nœuds avec 5,4 TB de mémoire [81]. Pour implanter la génomique des pathogènes, PHE a dû investir 2,0 et 1,8 millions \$US dans respectivement l'équipement de séquençage (« wetlab ») et d'analyse de données (« drylab ») (correspondant à 1,7 et 1,5 millions £) [81]. PHE a dû aussi remodeler les espaces des laboratoires, revoir les dispositions légales de la santé publique, remodeler ses procédures analytiques et former les équipes [81]. Contrairement à ses activités passées, les laboratoires de contrôle des pathogènes de PHE sont aujourd'hui organisés dans une unique pièce, utilisent une procédure unique et emploient des analystes et développeurs spécialisés en bioinformatique [81]. Ces laboratoires de PHE ont réduit ces dernières années l'utilisation d'animaux, le temps de manipulation des échantillons, ainsi que le coût de l'encadrement [81]. Ces laboratoires de PHE sont aujourd'hui davantage multidisciplinaires et certifiés ISO15189 pour le WGS [81].

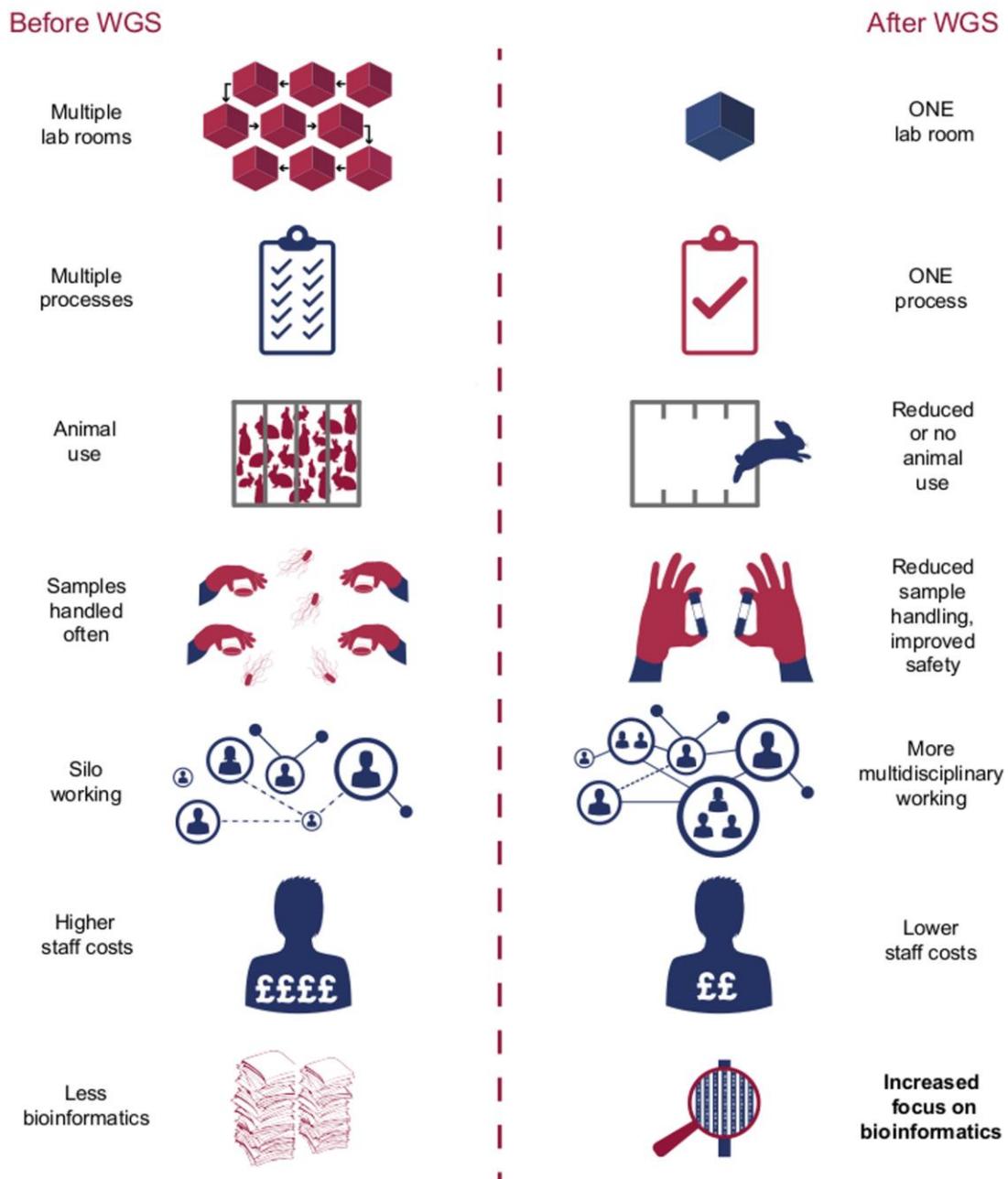


Figure 2 : Éléments modifiés dans l'organisation des laboratoires de « Public Health England » (PHE) avant et après l'implémentation du « whole genome sequencing » (WGS) des pathogènes alimentaires majeurs (*Salmonella*, *Escherichia coli*, *Shigella*, *Listeria* et *Campylobacter*)

III. Les enjeux du mémoire

Avec une description large des éléments m'amenant à candidater à l'HDR, le présent mémoire découle des instructions légales et institutionnelles (intro. § III.1) et est totalement personnel tant dans son organisation que de sa rédaction (intro. § III.2).

III.1. Instructions légales et institutionnelles

D'un point de vue légal, l'arrêté du 23 novembre 1988 [2], modifié par les arrêtés des 13 février 1992 [9], 13 juillet 1995 [5] et 25 avril 2002 [6], stipule que le dossier de candidature à l'HDR « *comprend soit un ou plusieurs ouvrages publiés ou dactylographiés, soit un dossier de travaux, accompagnés d'une synthèse de l'activité scientifique du candidat permettant de faire apparaître son expérience dans l'animation d'une recherche* ». D'un point de vue institutionnel, mon lieu d'accueil me propose un guide de rédaction de l'HDR, visant notamment à présenter son établissement et veiller à souligner les compétences du candidat en termes de publications, encadrement scientifique, élaboration et soumission de projets de recherche, participation à des réseaux scientifiques, reconnaissance internationale, animation scientifique, ainsi que capacité à élaborer un programme de recherche [82].

III.2. Point de vue personnel

En raison de mon parcours de chercheur en microbiologie, et plus particulièrement en génomique bactérienne dans les domaines de la santé humaine et chez l'animal, ainsi que de la sécurité environnementale et alimentaire, le présent mémoire est exposé sous forme d'une synthèse d'activités scientifiques supportées par un dossier de travaux. Plus précisément, le présent mémoire de candidature à l'HDR est proposé à l'université Paris-Est (UPE) avec le support de mon lieu d'accueil qui est un membre associé à l'université Paris-Est : l'agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (ANSES). Suite à la présentation du lieu d'accueil révélateur des tenants et aboutissants de mon parcours professionnel passé et futur (CHAPITRE I : LIEU D'ACCUEIL), un bilan de carrière sera exposé et supporté par des activités et travaux accomplis dans différents domaines de la microbiologie en constante mutation (CHAPITRE II : BILAN DE CARRIÈRE), afin d'annoncer un programme de recherche futur constitué de plusieurs projets faisant écho aux deux éléments précédents (CHAPITRE III : PROGRAMME DE RECHERCHE).

« L'art ne reproduit pas le visible, il le rend visible »

Paul Klee

CHAPITRE I : LIEU D'ACCUEIL

Dans le présent chapitre est exposé mon lieu de travail actuel (CHAPITRE I : LIEU D'ACCUEIL) qui soutient mon application à l'obtention de l'HDR en raison de mes expériences en recherche (Annexe 1), mon bilan de carrière en laboratoire (CHAPITRE II : BILAN DE CARRIÈRE) et mon programme de travaux de recherche pour les années à venir (CHAPITRE III : PROGRAMME DE RECHERCHE).

I. L'agence d'accueil : Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail

L'ANSES (Adresse web 1) est née le 1^{er} juillet 2010 de la fusion de deux agences sanitaires : l'agence française de sécurité sanitaire des aliments (AFSSA) et l'agence française de sécurité sanitaire de l'environnement et du travail (AFSSET). Son rayon d'activité (chap. I § I.1) et son organisation (chap. I § I.2) font de l'ANSES un établissement unique visant selon sa devise à « *Connaître, évaluer et protéger* » (Figure 3).

I.1. Rayon d'activités

Les tutelles (chap. I § I.1.a), missions (chap. I § I.1.b) et champs de compétences (chap. I § I.1.c) de l'ANSES font de cette agence un acteur majeur de la santé publique en France. L'ANSES met en œuvre des expertises scientifiques indépendantes et pluralistes sur la base de 1 350 agents, 800 experts extérieurs mobilisés, 142 millions \$ US de budget annuel (130 millions €), 8 000 avis émis depuis 1999, 80 mandats de référence nationaux, 250 publications scientifiques par an, ainsi que 100 doctorants et postdoctorants.

I.1.a. Tutelles

À sa création, l'ordonnance n°2010-18 du 7 janvier 2010 [83] fixe les missions de l'ANSES. Placée sous la tutelle des ministres chargés de la santé, de l'agriculture, de l'environnement, du travail, ainsi que de la consommation, l'ANSES est un établissement public à caractère administratif (EPA). L'ANSES vise à éclairer la politique sanitaire des pouvoirs publics via l'évaluation des risques dans le domaine de l'alimentation, de l'environnement et du travail.

I.1.b. Missions

En s'appuyant sur ses expertises scientifiques indépendantes et pluralistes, l'ANSES a pour principal objectif d'assurer la sécurité sanitaire humaine dans les domaines de l'environnement, du travail et de l'alimentation. L'ANSES intervient aussi dans la protection de la santé et du bien-être des animaux, la protection de la santé des végétaux, l'évaluation des propriétés nutritionnelles et fonctionnelles des aliments, ainsi que certains domaines liés aux médicaments vétérinaires. Dans ces champs de compétences, l'ANSES organise et finance des programmes de recherche scientifique et technique, tout en assurant ses missions de veille, alerte, vigilance et référence. L'ANSES recommande à ces autorités les mesures de police sanitaire à mettre en œuvre en cas de menaces par un danger grave. À la demande du gouvernement, l'ANSES représente la France en participant aux travaux des instances européennes et internationales.

I.1.c. Champs de compétences

L'ANSES a la possibilité d'établir des relations conventionnelles avec une trentaine d'établissements français, incluant des établissements d'enseignement et de recherche présentant des missions complémentaires (Annexe 2). L'ANSES est aussi en capacité de s'impliquer avec des personnes publiques ou privées, françaises ou étrangères, dans le cadre de consortium d'intérêt public (Annexe 2). Les implications de l'ANSES passent

par une surveillance (chap. I § I.1.c.i), une expertise (chap. I § I.1.c.ii) et une recherche (chap. I § I.1.c.iii), toutes les trois focalisées sur l'évaluation des risques sanitaires et nutritionnels.

1.1.c.i. Surveillance et vigilance

Ayant comme objectif de veiller sur différentes problématiques sanitaires, la surveillance menée par l'ANSES participe à la création de bases de données scientifiques et techniques. La surveillance de l'ANSES met aussi en place des observatoires sur les produits et procédés entrant dans son champ de compétences (chap. I § II.1.a). Entre autres domaines de compétences, l'ANSES est aussi impliquée dans des systèmes de vigilance : nutrivigilance, phytopharmacovigilance et toxicovigilance.

1.1.c.ii. Expertise

En s'appuyant sur des comités d'experts spécialisés (CES), l'ANSES organise une expertise collective répondant à ses propres autosaisines et saisines qui lui sont transmises. Les CES répondent à ces saisines et autosaisines liées à l'évaluation des risques sanitaires, sur la base des données existantes à travers différents dispositifs d'animation et de coordination de réseaux d'organismes publics. Les activités des laboratoires de l'ANSES sont partie prenante de l'expertise à l'ANSES, en exerçant des activités nationales et communautaires de référence. Un pôle majeur de l'expertise à l'ANSES est dédié aux appuis scientifique et technique nécessaires aux autorités compétentes pour l'évaluation des substances, en particulier des produits biocides et chimiques. Visant à délivrer des autorisations de mise sur le marché, l'ANSES couvre aussi l'évaluation des produits et adjuvants des molécules impliquées en phytopharmaceutiques, fertilisants et supports de culture.

1.1.c.iii. Recherche

Visant en premier lieu à contribuer à l'information, la formation et la diffusion scientifique et technique, la recherche à l'ANSES participe aux débats publics nationaux et communautaires en les questionnant et en répondant à ses interrogations. Les activités de recherche à l'ANSES sont organisées dans les laboratoires travaillant sur différentes problématiques. Ces activités de recherche sont menées dans le cadre de collaborations aux échelles nationales et communautaires.

Figure 3 : Logos français (à gauche) et anglais (à droite) de l'agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (ANSES)



1.2. Organisation

En raison de la diversité de ses missions (chap. I § I.1.b) et de la multidisciplinarité de ses champs de compétences (chap. I § I.1.c), l'ANSES répond à ses tutelles (chap. I § I.1.a) en organisant ses travaux et services sous des formes fonctionnelle (chap. I § I.2.a), thématique (chap. I § I.2.b) et hiérarchique (chap. I § I.2.c).

1.2.a. Fonctionnelle

L'ANSES est ouverte sur la société en organisant des comités de dialogue réguliers et spécifiques de thématiques données avec les parties prenantes, en particulier issues de la société civile. L'organisation fonctionnelle de l'ANSES se décline en un conseil administratif (chap. I § I.2.a.i), un conseil scientifique (chap. I § I.2.a.ii) et une expertise collective (chap. I § I.2.a.iii).

1.2.a.i. Conseil d'administration

En plus d'un président et des membres du personnel, le conseil d'administration de l'ANSES est composé de cinq collèges du grenelle de l'environnement. Ces collèges intègrent des représentants de l'état, de partenaires sociaux, d'organisations professionnelles, d'associations et d'élus. Les membres de ces collèges se partagent des droits de vote, avec une moitié de ces derniers pour le collège des représentants de l'état. S'ajoute à ce conseil d'administration des personnalités qualifiées et les représentants du personnel de l'ANSES. Ce conseil administratif compte aussi des comités d'orientation thématiques constitués de personnalités extérieures ayant des visions sur les tendances de la société civile (santé-environnement, santé-travail, alimentation, santé et bien-être animal, santé végétale). Le conseil d'administration supervise le processus d'évaluation de l'activité de recherche de l'ANSES.

1.2.a.ii. Conseil scientifique

Indépendant et exclusivement composé de scientifiques, le conseil scientifique est présidé en donnant une large place aux scientifiques, notamment étrangers. Ce conseil scientifique compte trente membres et garantit la qualité scientifique de l'expertise, ainsi que son indépendance. Dans un objectif de transparence, les membres du conseil scientifique, tout comme l'ensemble du personnel permanent de l'ANSES, sont soumis à une déclaration publique d'intérêt (DPI). Cette DPI vise à prévenir tous conflits d'intérêt entre les agents de l'ANSES et leurs collaborateurs extérieurs.

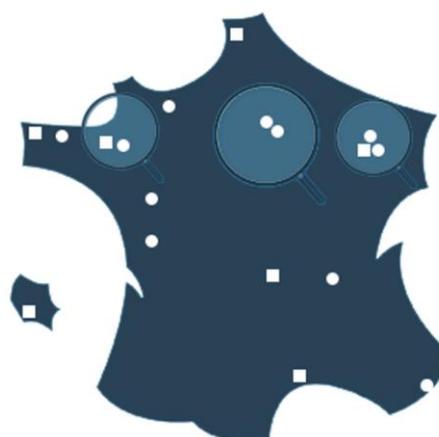
1.2.a.iii. Experts

L'expertise collective mobilise 800 experts extérieurs et permet à l'ANSES de fournir des avis et/ou recommandations rendus systématiquement publics en prenant en compte l'ensemble des données scientifiques disponibles et des opinions exprimées. Dans le cadre de cette expertise collective au plus près des filières, un comité de déontologie et de prévention des conflits d'intérêt peut intervenir dans toutes les situations. Il peut être saisi par un membre du conseil d'administration, du conseil scientifique, des comités d'experts spécialisés, le directeur général ou un agent de l'ANSES.

1.2.a.iv. Laboratoires

Les 9 laboratoires de référence et de recherche de l'ANSES sont répartis sur l'ensemble du territoire (Figure 4). Au plus près des filières, ces laboratoires traitent des domaines « santé et bien-être des animaux », « sécurité sanitaire des aliments » et « santé des végétaux ». Avec un rôle majeur dans la connaissance des dangers, ces laboratoires supportent les activités de l'expertise collective. Ils ont pour objectifs d'expertiser, surveiller épidémiologiquement, alerter, ainsi qu'assister scientifiquement et techniquement. En se basant sur des expertises en épidémiologie, microbiologie et résistance aux antimicrobiens, toxines et contaminants physicochimiques, les laboratoires de l'ANSES collectent des données issues des réseaux de laboratoires agréés.

Figure 4 : Laboratoires (cercles) et antennes de laboratoires (carrés) de l'agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (ANSES) à Angers, Boulogne-sur-Mer, Clermont-Ferrand, Dozulé, Fougères, Maisons-Alfort, Malzéville, Montpellier, Nancy, Niort, Ploufragan, Plouzané, Rennes, Saint-Pierre de la Réunion et Sophia-Antipolis



1.2.b. Thématique

Les activités des 7 thématiques de l'ANSES sont validées par le conseil d'administration, le conseil scientifique et l'expertise collective (chap. I § 1.2.a). À ce jour, les thématiques de l'ANSES sont « alimentation et nutrition humaine », « santé-travail, santé environnement », « alimentation et santé animale », « médicament vétérinaire - agence nationale du médicament vétérinaire (ANMV) », « santé et protection du végétal », « produits phytopharmaceutiques », ainsi que « biocides et fertilisants ». Ces 7 thématiques sont gérées par le système de management de la qualité de l'ANSES qui s'appuie sur la norme ISO 9001 pour intégrer les activités de réalisation, de support et de management (c.à.d. NF X 50-110, NF EN ISO/CEI 17025, NF EN ISO/CEI 17043 et NF EN ISO/CEI 17020). Tout comme l'ensemble de l'ANSES, son système de management de la qualité s'engage à mettre en œuvre les dispositions de la « charte du développement durable des établissements publics et entreprises publiques » [84] et de la « charte de l'ouverture à la société » [85]. Reposant sur 3 piliers, à savoir « *connaître, évaluer et protéger* », la politique qualité de l'ANSES s'engage dans une démarche d'excellence scientifique, de réactivité face aux situations de crises et d'anticipation des dangers émergents. Elle assure son indépendance, sa transparence et son ouverture à la société.

1.2.c. Hiérarchique

Les organisations fonctionnelle (chap. I § 1.2.a) et thématique (chap. I § 1.2.b) de l'ANSES sont instituées d'un point de vue hiérarchique sous le directeur général Roger Genet. Cette direction générale est soutenue par une direction des affaires européennes et internationales et une direction de la communication et des relations institutionnelles (Annexe 3). Ces directions chapotent 4 pôles nommés « affaires générales », « produits réglementés », « sciences pour l'expertise », ainsi que « recherche et référence ». Le pôle « affaires générales » assure la responsabilité hiérarchique des directions des ressources humaines, des achats, de l'administration et des finances (DAF), technique et informatique (DTI), des affaires juridiques, et de la qualité et de l'audit interne. Ce pôle « affaires générales » intègre aussi une mission hygiène, sécurité, sûreté. Le pôle « produits réglementés » assure la cohérence d'ensemble des processus d'évaluation des produits réglementés. Le pôle « produits réglementés » a la responsabilité de la délivrance des autorisations de mise sur le marché et de vigilances, en particulier les médicaments vétérinaires mis en œuvre par l'ANMV. Le pôle « sciences pour l'expertise » assure la préparation de réunions du conseil scientifique de l'agence avec son président. Ce pôle « sciences pour l'expertise » coordonne les activités de la direction de l'évaluation des risques (DER), de la mission sciences sociales, expertise et société, de la direction alerte et veille sanitaires et de la direction financement de la recherche et veille scientifique. Le pôle « recherche et référence » comporte une direction de la stratégie et des programmes (DSP) chapotant les 9 laboratoires distribués sur le sol français (Figure 4).

II. Laboratoire d'accueil : Laboratoire de sécurité des aliments

Parmi les laboratoires de l'ANSES (chap. I § 1.2.c), le laboratoire de sécurité des aliments (LSAL) est sous la direction de Laurent Laloux et de son adjointe Anne Brisabois. Le LSAL est organisé sur 2 sites : Maisons-Alfort (MA) et Boulogne-sur-Mer (BsM) (Figure 4). Les activités et travaux associés à la sécurité sanitaire des aliments sont pris en charge par 120 personnes sur le site de Maisons-Alfort. Les activités et travaux concernant la qualité et la sécurité des produits de la pêche et de l'aquaculture sont gérées par plus de 20 personnes à Boulogne-sur-Mer. L'organisation des missions du LSAL (chap. I § II.2.a) se décline au regard d'une segmentation par mandats de référence (chap. I § II.2.b). Ces mandats de référence sont par ailleurs intégrés aux activités de surveillance (chap. I § II.1.a), référence (chap. I § II.1.b) et recherche (chap. I § II.1.c) du LSAL.

II.1. Rayon d'activités

Les mandats nationaux et européens du LSAL (chap. I § II.2.b) se déclinent en des activités de surveillance (chap. I § II.1.a) et référence (chap. I § II.1.b). Elles-mêmes sont nourries et soutenues par des travaux de recherche dans le cadre de projets propres, le plus souvent conventionnés (chap. I § II.1.c) (Figure 5).

II.1.a. Surveillance

Les activités de surveillance au LSAL datent de plusieurs années, notamment de par son implication dans le réseau de surveillance des isolats de *Salmonella* d'origine non humaine et sa participation aux plans de surveillance pilotés par la direction générale de l'alimentation (DGAL). Actuellement, le LSAL est impliqué fortement dans une plateforme de surveillance sanitaire de la chaîne alimentaire (SCA) créée en juillet 2018. Sous l'impulsion des états généraux de l'alimentation soulignant la nécessité de renforcer le dispositif national de surveillance et de prévention des risques sanitaires, cette plateforme SCA est une action collaborative [86]. Elle mobilise les acteurs de la chaîne alimentaire pour assurer la sécurité des consommateurs en étroites synergies avec les plateformes d'épidémiologie en santé végétale (ESV) et animale (ESA) (Annexe 4). Cette plateforme SCA vise à implémenter un système de surveillance de tous les contaminants des aliments en valorisant les plans de surveillance, plans de contrôle et autocontrôles réalisés par les industriels. Néanmoins, des nécessités d'actions prioritaires sur *Salmonella* et *Campylobacter* ont été émises en raison des importantes prévalences dans l'alimentation et incidences chez l'homme de ces bactéries pathogènes [87]. À l'échelle européenne, ces pathogènes sont effectivement les plus fréquemment transmis à l'homme par des aliments, ainsi que les plus fréquemment associés à des zoonoses alimentaires en 2015 [88], 2016 [89] et 2017 [87]. Les activités de surveillance du LSAL collaborent avec l'activité de recherche (chap. I § II.1.c).

II.1.b. Référence

Les activités de référence visent à soutenir les mandats nationaux et européens du LSAL en apportant aux administrations de tutelle des expertises scientifiques et techniques. Les laboratoires de référence de l'union européenne (LRUE) ont principalement des échanges internationaux avec les laboratoires nationaux de référence (LNR) correspondants. Ils animent le réseau de LNR ainsi constitué et les LNR communiquent à l'échelle nationale avec les laboratoires de leurs domaines d'activités. En collaborant avec l'activité de recherche (chap. I § II.1.c), l'activité de référence du LSAL est donc notamment dédiée à la collecte d'échantillons et de données transmises par les réseaux de laboratoires agréés à l'échelle nationale et de LNR à l'échelle européenne. L'activité de référence réalise aussi des analyses, plans de surveillance, développements et validations de méthodes analytiques, tout en contribuant au transfert méthodologique par la formation et l'organisation des essais interlaboratoires d'aptitude. L'activité de référence promulgue aussi des formations agréées et reconnues. Par ailleurs, les activités de référence des LRUE du LSAL travaillent en étroites relations avec les autorités européennes : l'EFSA et l'ECDC.

II.1.c. Recherche

Les activités de recherche du LSAL sont supportées par des directeurs de thèse avec HDR, chargés de projets scientifiques, postdoctorants, doctorants et étudiants en master. Le LSAL est labellisé « équipe d'accueil » pour les écoles doctorales (ED) agriculture, alimentation, biologie, environnement, santé (ABIES, n°0581) et sciences de la matière, du rayonnement et de l'environnement (SMRE, n°0104), respectivement pour les sites de Maisons-Alfort et Boulogne-sur-Mer. L'activité de recherche du LSAL vise principalement à surveiller des dangers relatifs aux molécules et agents pathogènes de l'alimentation. Pour se faire, l'activité de recherche du LSAL se nourrit des données récoltées par les activités de surveillance (chap. I § II.1.a) et référence (chap. I § II.1.b) pour développer des méthodes analytiques novatrices. Par la suite, cette activité de

recherche promulgue des formations à l'utilisation des méthodes analytiques qu'elle élargit aux activités de surveillance (chap. I § II.1.a) et référence (chap. I § II.1.b) du LSAL. Décrire, comprendre et modéliser les comportements de ces molécules et agents pathogènes de l'alimentation font partie intégrante des travaux de recherche du LSAL. Les projets de recherche en question sont à la fois financés par des appels d'offres européens ou nationaux. Concernant les appels d'offres européens, doivent être cités « seventh framework programme » (FP7), « horizon 2020 » (H2020), « collaborative management platform for detection and analyses of re-emerging and foodborne outbreaks in Europe » (COMPARE) et « European joint programme » (EJP). Concernant les appels d'offres nationaux, doivent être cités l'agence nationale de la recherche (ANR), les actions de recherche concertées d'initiative régionale (ARCIR), les contrats de plan État-Région (CPER), le fonds unique interministériel (FUI), les domaines d'intérêt majeur franciliens (DIM) et l'établissement national des produits de l'agriculture et de la mer (FranceAgriMer).

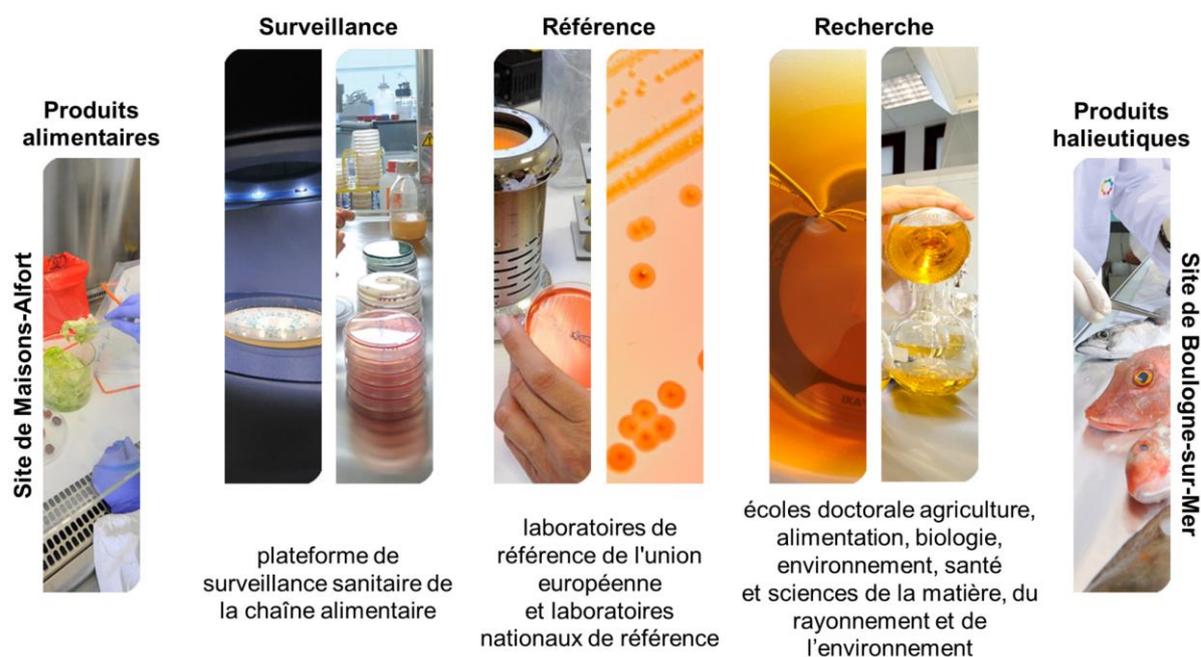


Figure 5 : Domaines d'activité du laboratoire de sécurité des aliments (LSAL) sur les sites de Maisons-Alfort et Boulogne-sur-Mer

II.2. Organisation

Les missions du LSAL sont orientées vers la production de données et l'identification des dangers (pathogènes ou contaminants chimiques), ainsi que l'évaluation des risques moléculaire et biologique liés à l'alimentation (chap. I § II.2.a). Les mandats correspondants à ces activités sont organisés par départements, unités et missions transversales (chap. I § II.2.b).

II.2.a. Missions

Avec comme principale mission d'identifier les dangers dans les produits alimentaires, le LSAL développe et normalise des méthodes de détection, caractérisation et quantification, tout en participant à l'élaboration de leurs réglementations. Le LSAL analyse de plus les causes d'apparition de ces dangers et leurs facteurs de développement, en participant aux évaluations quantitatives des risques et estimations des prévalences des dangers de molécules et agents pathogènes. Ce laboratoire a aussi pour missions de surveiller et signaler l'émergence et la réémergence de ces agents pathogènes, en prenant en compte dans les analyses des risques, l'influence des pratiques et procédés des filières alimentaires. Ces activités du LSAL visent à améliorer

la maîtrise ces agents pathogènes tout au long de la durée de vie du produit alimentaire, usuellement désignée « *de la fourche à la fourchette* ».

II.2.b. Mandats

Intégrant respectivement 10 et 2 mandats nationaux et européens, le LSAL est organisé en 3 départements, 9 unités et 3 missions transversales. Effectivement, les activités des LNR et LRUE sont intégrées aux différentes unités et missions du LSAL (Annexe 5).

III. Équipe d'accueil : Mission génome analyse modélisation risque

La mission génome analyse modélisation risque (GAMeR) a été mis en place à partir de 2014 et officialisée en 2015 (Figure 6) par le chef du département contaminants microbiologiques des aliments, Michel-Yves Mistou. Il a pris la responsabilité de cette mission GAMeR pour implanter la génomique au LSAL (Annexe 5) avec le soutien des directions de l'agence (chap. I § III.3). L'organisation (chap. I § III.1) et les objectifs (chap. I § III.2) de cette mission GAMeR sont transversaux à toutes les unités présentant des activités en microbiologie au LSAL. J'ai personnellement intégré le 11 mai 2015 cette mission GAMeR dans laquelle je travaille toujours à l'heure actuelle en tant que chargé de projets scientifiques conduisant des travaux de recherche en génomique bactérienne.

Figure 6 : Logos français (à gauche) et anglais (à droite) de la mission genome analyse modélisation risque (GAMeR)



III.1. Organisation

Au sein du LSAL, la mission GAMeR est donc transversale aux unités impliquées en microbiologie des départements contaminants microbiologiques des aliments et produits de la pêche et de l'aquaculture (Annexe 6). En 6 ans d'existence (non officiellement : 2014-2015 et officiellement : 2016-2019), la mission GAMeR a comporté en plus de son responsable : 1, 2 et 3 chargés de projets scientifiques en contrat à durée indéterminée (CDI), respectivement depuis 2014, 2015 et 2018. Depuis le début de son existence, la mission GAMeR a aussi intégré annuellement entre 2 et 4 étudiants en master, doctorat ou postdoctorat. Actuellement en 2019, la mission GAMeR comporte 1 responsable, 3 chargés de projets scientifiques en CDI, 1 postdoctorant, 2 doctorantes et 1 étudiante de master en cursus d'apprentissage. La mission GAMeR intègre des profils d'analystes et développeurs en bioinformatique, tous les deux tournés vers la gestion de données omiques massives. Les publications scientifiques représentent une valorisation essentielle des activités de la mission GAMeR.

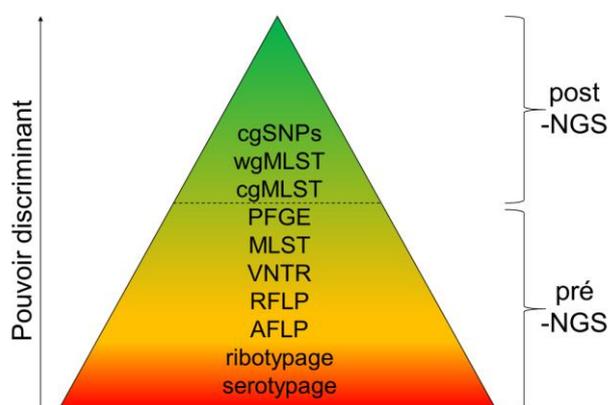
III.2. Objectifs

L'OMS définit le NGS comme un outil d'intérêt pour la surveillance et la réponse aux évènements de toxi-infections alimentaires collectives (TIAC) [90] et il existe une forte compétition nationale et internationale dans le secteur de la génomique bactérienne (Tableau 1). Dans ce contexte, les activités de recherche de la mission GAMeR se sont construites sur la base de compétences en bioinformatique, statistique et gestion de réseau informatique. La mission GAMeR promulgue des expertises en analyse des données omiques et modélisation du risque microbiologique dans les aliments. Les objectifs de la mission GAMeR intègrent du typage (chap. I § III.2.a), de l'attribution des sources (chap. I § III.2.b), de la recherche avancée en génomique (chap. I § III.2.c) et de la diffusion de l'expertise (chap. I § III.2.d). Ces objectifs n'ont pas d'ordres de priorité et sont donc établis en fonction des urgences à traiter.

III.2.a. Typage

Le LSAL a historiquement développé et appliqué des méthodes de caractérisation bactérienne au fur et à mesure des développements analytiques par typage : sérotypage, ribotypage, « amplified fragment-length polymorphism » (AFLP), « restriction fragment length polymorphism » (RFLP), « variable number tandem repeat » (VNTR), « multilocus sequence typing » (MLST) et « pulsed-field gel electrophoresis » (PFGE). Les objectifs de ces méthodes de typage sont d'investiguer des épisodes de TIAC et de mieux connaître les clones circulants dans les élevages et les industries agroalimentaires (IAA). Même si ces développements ont apporté des méthodes de caractérisation de plus en plus résolutes (Figure 7), ces dernières sont longues à mettre en œuvre (de quelques jours à plusieurs semaines) et ne sont pas systématiquement applicables à tous les genres bactériens. À l'instar de ces méthodes conventionnelles de caractérisation, l'avènement récent du NGS (intro. § II.1) et l'amélioration des méthodes analytiques dans ce domaine (intro. § II.3) ont apporté des solutions analytiques. Ces solutions analytiques visent à caractériser les bactéries sur l'ensemble de leurs génomes indépendamment du genre considéré, tout en diminuant les coûts (intro. § II.2) et les délais analytiques (intro. § II.4). À titre d'exemple concernant ces méthodes discriminantes basées sur le NGS, les approches « coregenome MLST » (cgMLST), « whole genome MLST » (wgMLST) et « coregenome SNPs » (cgSNPs) doivent être soulignées (Figure 7). En se basant sur une veille des technologies NGS, la mission GAMeR s'investit dans le développement de méthodes de typage très résolutes des microorganismes pathogènes issus de l'alimentation sur la base de l'analyse de génomes. Par le biais de ses travaux, la mission GAMeR contribue à transférer des compétences dans les unités de microbiologie impliquées dans la mise en œuvre des activités de surveillance et référence du LSAL [91, 92].

Figure 7 : Pouvoir discriminant de méthodes de caractérisation pré- et post-next generation sequencing (NGS) : sérotypage, ribotypage, « amplified fragment-length polymorphism » (AFLP), « restriction fragment length polymorphism » (RFLP), « variable number tandem repeat » (VNTR), « multilocus sequence typing » (MLST), « pulsed-field gel electrophoresis » (PFGE), « coregenome MLST » (cgMLST), « whole genome MLST » (wgMLST) et « coregenome SNPs » (cgSNPs)



III.2.b. Attribution des sources

En étroite relation avec la DER en charge de l'évaluation des risques, la mission GAMeR participe à faire évoluer des modèles d'attribution des sources et de dose-réponse en y intégrant des données omiques [93–95]. Ceci a pour but final d'apprécier l'exposition des consommateurs aux dangers microbiologiques considérés.

III.2.c. Recherche avancée en génomique

Tout en définissant son propre programme de recherche, la mission GAMeR développe une recherche en collaboration avec les équipes concernées du LSAL dans les domaines de la génomique et transcriptomique des bactéries pathogènes transmises par les aliments (bactéries pathogènes alimentaires) [96–99, 99–104].

III.2.d. Diffusion de l'expertise

Que ce soit pour les activités des LNR ou LRUE du LSAL, la mission GAMeR promulgue des formations théorique et pratique annuelles à l'utilisation du réseau Linux qu'elle a développés depuis 2014 et des outils stables implémentés dans ce dernier. À l'échelle nationale, la mission GAMeR transfère ses développements stables via des systèmes d'archivage en ligne à la plateforme génomique de Ploufragan. En collaboration avec le laboratoire de santé animale (LSAN), la mission GAMeR s'implique aussi dans l'organisation trimestrielle d'actions d'animation scientifique via le groupe débats autour de la bioinformatique et la génomique (DEBUG) sur le site de Maisons-Alfort.

III.3. Supports

En étroite relation avec la DSP, la mission GAMeR initie et poursuit la mise en place de marchés publics portant sur du NGS dans le cadre de l'acquisition de données par le LSAL (Annexe 3). En parallèle de ses développements d'outils analytiques, la mission GAMeR initie et conserve des relations de confiance avec la DTI en vue de faire évoluer le réseau Linux de traitement de données qu'elle développe.

Tableau 1 : Initiatives nationale et internationale de plateformes analytiques en génomique des microorganismes pathogènes et autres microorganismes d'intérêt isolés chez l'Homme, d'animaux, de l'alimentation humaine et animale, ainsi que d'environnement de production alimentaire

Plateforme	Signification de l'acronyme	Organisation	Adresse web	Référence
BIGSdb	bacterial isolate genome sequence database	Université d'Oxford	https://bigsdbs.readthedocs.io/en/latest/	[105]
BIGSdb-Lm	bacterial isolate genome sequence database for <i>Listeria monocytogenes</i>	Institut Pasteur	https://bigsdbs.pasteur.fr/listeria/	[106]
BioNumerics	Sans objet	Applied Maths (Biomerieux)	http://www.applied-maths.com/bionumerics	[80]
CGE	center for genomic epidemiology	Université technique danoise	http://www.genomicepidemiology.org/	[80]
COMPARE	collaborative management platform for detection and analyses of re-emerging and foodborne outbreaks in Europe	Institut européenne de bioinformatique et université technique danoise	https://www.compare-europe.eu/	[80]
Enterobase	<i>Enterobacteriaceae</i> database	Université de Warwick	https://enterobase.warwick.ac.uk/	[107]
GenomeTrakr	Sans objet	US food and drug administration	https://www.fda.gov/food/whole-genome-sequencing-wgs-program/genometrakr-network	[108]
INNUENDO	a cross-sectorial platform for the integration of genomics in surveillance of food-borne pathogens	Université de Lisbonne	https://innuendo.readthedocs.io/en/latest/	[109]
IRIDA	integrated rapid infectious disease analysis	Agence de santé public du Canada	https://www.irida.ca/	[110]
Microreact	Sans objet	Public health England (PHE), CDC, Microbiology Society, Institut Norvégienne de santé Publique (NIPH) et Wellcome Trust	https://microreact.org/showcase	[111]
Pathogen Watch	Sans objet	Institut Wellcome Sanger	https://pathogen.watch/	[80]
PubMLST	public databases for multilocus sequence typing	Université d'Oxford et Wellcome Trust	https://pubmlst.org/contact.shtml	[112]
PulseNet	Sans objet	Centre de contrôle et prévention des maladies (CDC)	https://www.cdc.gov/pulsenet/index.html	[113]
SeqSphere	Sans objet	Ridom GmbH	https://www.ridom.de/seqsphere/	[114]

« L'empirisme n'est point la négation de la science expérimentale [...], ce n'en est que le premier état »

Claude Bernard

CHAPITRE II : BILAN DE CARRIÈRE

Le présent bilan de carrière (CHAPITRE II : BILAN DE CARRIÈRE) fait référence aux recherches que j'ai entrepris dans mon lieu d'accueil actuel, mais aussi à mes activités antérieures (Annexe 1 et Adresse web 2). Toutes les deux introduisent un programme de recherche qui sera présenté en dernier lieu (CHAPITRE III : PROGRAMME DE RECHERCHE). Le présent bilan de carrière est exposé au regard de 27 travaux, constitués de mon manuscrit de thèse obtenue le 28 février 2011 [115], des demandes de brevets soumisses en 2018 [91], un chapitre d'ouvrage publié [116], 2 publications internationales dans lesquelles mes contributions sont remerciées [96, 117], ainsi que 20 publications internationales publiées en tant que co-auteur ou auteur principal [92, 94, 95, 97–102, 118–128] et 2 en cours de soumission pour publication [103, 104]. À l'exclusion du manuscrit de thèse [115], des demandes de brevets [91], du chapitre d'ouvrage [116] et des remerciements de participation [96, 117], ces publications internationales représentent mes contributions en tant que dernier auteur [101, 102], premier auteur [92, 118, 120, 122–127] et auteur intermédiaire [94, 95, 97–100, 103, 104, 119, 121, 128]. L'ensemble de ces publications internationales sont accessibles en ligne (Adresse web 3) et reflètent mes travaux de master [117, 127], thèse [116, 122, 124–126] et postdoctorat [118–121, 123], ainsi que mes activités de recherche actuelles [92, 94–104, 128]. Mon identifiant ouvert pour chercheur et contributeur (ORCID : « open researcher and contributor identifier ») est le 0000-0002-7480-4197 (Adresse web 4).

I. Contextes et financements des travaux

Mes présents travaux de recherche ont eu lieu dans les domaines de la santé humaine (chap. II § I.1) et animale (chap. II § I.2), ainsi que dans la sécurité environnementale (chap. II § I.3) et alimentaire (chap. II § I.4) (Tableau 2). Avec comme fil d'Ariane le développement et l'exploitation des méthodes analytiques en microbiologie, mes travaux ont contribué à faire évoluer ces domaines d'application de la génomique en microbiologie (chap. II § II).

I.1. Santé humaine

Conduit par la résurgence de la tuberculose dans les communautés Inuits du Nunavik entre 2001 et 2013, j'ai eu pour objectifs de développer en collaboration avec l'agence de la santé publique du Canada (PHAC), des solutions de compréhension de ce phénomène épidémique. Ces solutions de compréhension de ce phénomène épidémique ont par la suite été fournies à la régie régionale de la santé et des services sociaux du Nunavik (RRSSSN). Ces travaux majeurs en santé humaine ont eu lieu pendant mes années de postdoctorat à l'institut de recherche du centre de santé de l'université de McGill (RI MUHC). D'une part, ces travaux se sont plus précisément focalisés sur l'amélioration du diagnostic moléculaire de la maladie de Crohn [123]. D'autre part, ces travaux ont aussi porté sur la compréhension moléculaire [121] et génomique [119] de l'émergence de la bactérie intracellulaire stricte *Mycobacterium tuberculosis*. Plus précisément, mes travaux ont porté sur le décryptage génomique d'épidémies de tuberculose aux échelles d'un village [120] et d'une communauté [118] Inuits au Canada. En plus de mes contributions analytique et rédactionnelle, j'ai participé aux conceptions de ces travaux et aux recherches de financement correspondantes auprès du RI MUHC (grant #29836) et du fonds de recherche en santé Québec (FRQS) (grant #26274).

Tableau 2 : Domaines de la microbiologie (santé humaine et animale, sécurité environnementale et alimentaire), cursus, lieux d'accueil (établissement, laboratoire, équipe), sujets de recherche et sources de financement

Domaine	Cursus	Lieu d'accueil	Sujet de recherche	Source de financement
Sécurité alimentaire (2015-19)	Chargé de projet recherche en génomique bactérienne	ANSES LSAL GAMeR	Implémentation de la génomique bactérienne (<i>Salmonella</i> , <i>Listeria</i> , <i>Clostridium</i> , <i>Bacillus</i> , <i>Staphylococcus</i>) et évolution de l'adaptation aux hôtes de <i>Salmonella</i> et des plasmides bactériens	
Santé humaine (2011-15)	Postdoctorat en génomique bactérienne	McGill RI MUHC Behr Lab	Diagnostic moléculaire de la maladie de Crohn (<i>Mycobacterium avium</i>) et génomique de l'émergence de la tuberculose (<i>Mycobacterium tuberculosis</i>) dans des communautés Inuits au Canada	
Sécurité environnementale (2007-11)	Doctorat en microbiologie moléculaire et génomique	UPE ENPC LEESU	Développements de méthodes de détection et quantification en bactériologie, biologie moléculaire et génomique pour identifier les sources des mycobactéries non-tuberculeuses dans les bassins versants	
Santé animale (2004-06)	Master en microbiologie moléculaire	ANSES LSAL UZB	Phylogénie bactérienne et bases de la transmission des sous-espèces de <i>Mycobacterium avium</i> entre l'Homme, les porcs, les bovins, les oiseaux, ainsi que les faunes sauvage et domestique	

1.2. Santé animale

Mes travaux en santé animale ont eu lieu dans un contexte de suspicion de tuberculose bovine dans des élevages et de méconnaissance des bases de la transmission de la bactérie pathogène environnementale *Mycobacterium avium* entre l'Homme et les animaux. J'ai eu pour objectifs d'identifier les sous-espèces de *M. avium* responsables d'infections chez l'Homme, les porcs, les bovins, les oiseaux, ainsi que les faunes sauvage et domestique. Ces premiers travaux en santé animale datent de mon travail de master dans l'unité de zoonoses bactériennes (UZB) du LSAN, durant lequel j'ai étudié la phylogénie bactérienne et l'adaptation aux animaux de *M. avium* [127]. En plus de mes contributions analytique et rédactionnelle, j'ai participé aux conceptions de ces travaux et aux recherches de financement correspondantes auprès du « veterinary network of laboratories researching into improved diagnosis and epidemiology of mycobacterial diseases » (VENoMYC) en collaboration avec l'institut national de la recherche agronomique (INRA) (grant AFSSA-INRA #AIP P00297), dernièrement renommé institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE).

1.3. Sécurité environnementale

Mes travaux en sécurité environnementale ont été réalisés selon les directives européennes établissant des objectifs de préservation et restauration de l'eau de surface et de l'eau souterraine pour 2015 (2000/60/CE) et d'évaluation et de contrôle des effluents (91/271/CE et 98/15/CE). J'ai eu pour objectifs de développer et fournir aux autorités hospitalières et de contrôle de l'eau à Paris des outils analytiques pour prédire les comportements des mycobactéries non-tuberculeuses (MNT) dans les bassins versants. Ces autorités étaient plus précisément l'assistance publique hôpitaux de Paris (APHP), Eau de Paris et le syndicat interdépartemental pour l'assainissement de l'agglomération parisienne (SIAAP). Ces travaux font référence à mes activités en doctorat au laboratoire eau environnement et systèmes urbains (LEESU). Mes développements de méthodes de détection et quantification des MNT en bactériologie [126], biologie moléculaire [125] et génomique [122], m'ont permis d'identifier leurs sources majeures dans les bassins versants [115], à savoir l'eau de surface traitée par les stations d'épurations [124]. Ces travaux m'ont aussi permis de synthétiser les méthodes analytiques à mettre en œuvre dans le cadre de l'identification des sources de MNT dans l'eau de l'environnement [116]. Via ma participation à la conception de ces travaux, les financements correspondants ont été obtenus auprès du programme interdisciplinaire de recherche sur l'eau et l'environnement du bassin de la Seine (PIREN-Seine), de l'observatoire des polluants urbains (OPUR) et de la ville de Paris (grant #CRECEP-STEAs).

1.4. Sécurité alimentaire

En raison des bouleversements technologiques engendrés par l'avènement du NGS (intro. § II.1), j'ai comme objectif depuis 2015, d'implémenter la génomique bactérienne au LSAN de l'ANSES. J'occupe depuis cette date un poste de chargé de projets de recherche en génomique bactérienne dans la mission GAMeR (chap. I § III.2). À la fois au début de mes études supérieures (Annexe 1) et dans le cadre de mon poste actuel en génomique bactérienne (chap. I § III.2), j'ai été amené à travailler sur différentes bactéries pathogènes transmises par les aliments. Plus précisément, j'ai focalisé mes activités en génomique sur *Salmonella* [91, 92, 95, 96, 98, 100–102], *Listeria* [94, 99, 117, 128], *Clostridium* [97] et *Bacillus* [104], ainsi que sur les plasmides bactériens [103], en tant que collaborateur ou organisateur principal [92, 101, 102]. En plus des projets portés par des collaborateurs (chap. II § II.1), j'ai participé aux conceptions et obtentions de financements des projets portés par la mission GAMeR (chap. II § II.2). Ces projets ont été financés par le programme COMPARE (Horizon 2020, grant #643476), une collaboration nationale ANSES - INRA (grant #Typautobac) et le

projet « risk and disease burden of antimicrobial resistance » (RaDAR) du programme « One Health » EJP (grant #773830).

II. Encadrement et transversalité des travaux

Que ce soit dans les domaines de la santé humaine et animale, ou de la sécurité environnementale et alimentaire (chap. II § I), mes travaux m'ont conduit à encadrer des étudiants de master et doctorat. Ces travaux m'ont aussi conduit à échanger avec des collègues, dans le cadre de projets collaboratifs pilotés par ces derniers (chap. II § II.1) ou de projets sous ma responsabilité au sein de la mission GAMeR (chap. II § II.2).

II.1. Participation à des projets collaboratifs

J'ai eu des activités collaboratives concernant des projets portés par des collègues, depuis mes premières activités de jeune chercheur jusqu'à aujourd'hui. En effet, j'ai eu des activités de développements liés aux nouvelles technologies NGS (chap. II § II.1.a), j'ai contribué à la création d'un réseau Linux (chap. II § II.1.b) et j'ai transféré ces développements vers des étudiants et collègues (chap. II § II.1.c).

II.1.a. Amélioration continue des développements

Depuis mes premiers développements analytiques en génomique bactérienne visant à identifier des gènes caractéristiques des MNT [122], je n'ai eu de cesse d'encadrer des étudiants de différents niveaux (Annexe 1). J'ai aussi développé des outils dans ce domaine, tout en les améliorant au cours du temps, tant sur les plans de la justesse, que de la facilité d'exécution, afin de les transmettre à d'autres étudiants ou des collègues dans le cadre de projets collaboratifs. Les exemples les plus significatifs sont le développement et le transfert d'un outil d'identification des variants du coregénom (iVARcall2 : independent variant calling analysis) (Figure 8) et la mise en place d'un outil d'assemblage *de novo* de génomes bactériens (ARTwork : assembly of reads and typing workflow) (Figure 9). Ces outils iVARcall2 et ARTwork ont été adoptés par le LSAL depuis 2016 pour traiter environ 1 000 génomes annuellement. Les premières versions de mes développements de l'outil iVARcall2 datent de mes années de postdoctorat en 2012, coïncidant avec mon encadrement d'une étudiante en doctorat travaillant sur *M. tuberculosis* au RI MUHC (Robyn S. Lee : Annexe 1) [118, 120].

Figure 8 : Outil « independent variant calling analysis » (iVARcall2) développé par la mission génome analyse modélisation risque (GAMeR) au sein du laboratoire de sécurité des aliments (LSAL) pour identifier les variants de coregénomés bactériens par « mapping » et « variant calling » [96, 97, 99, 101, 102, 104, 118, 120, 128]

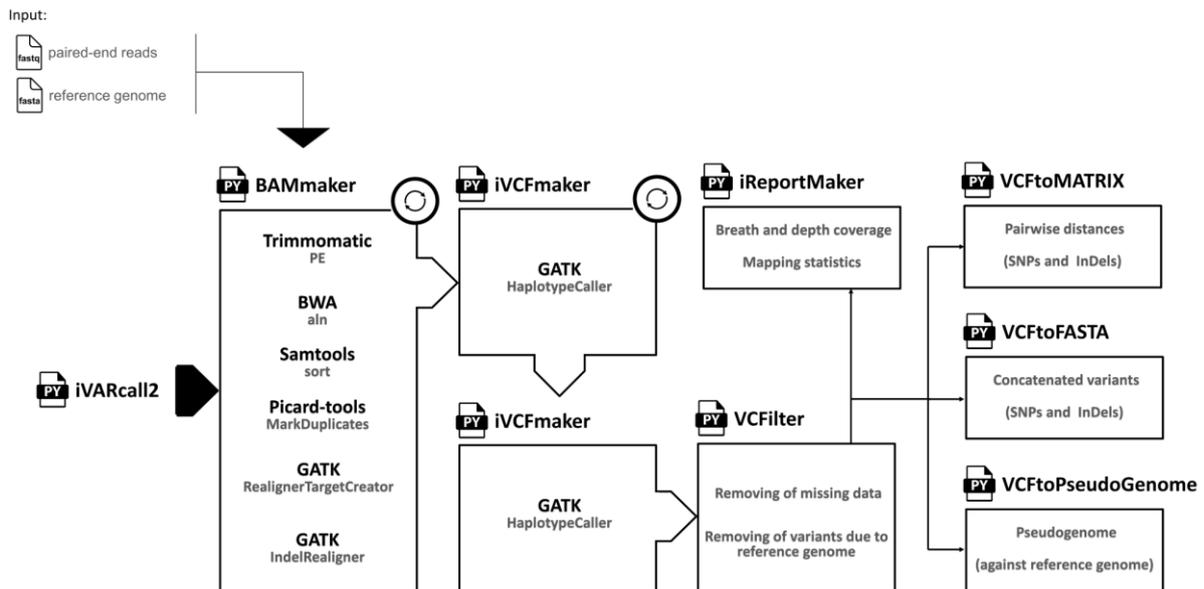
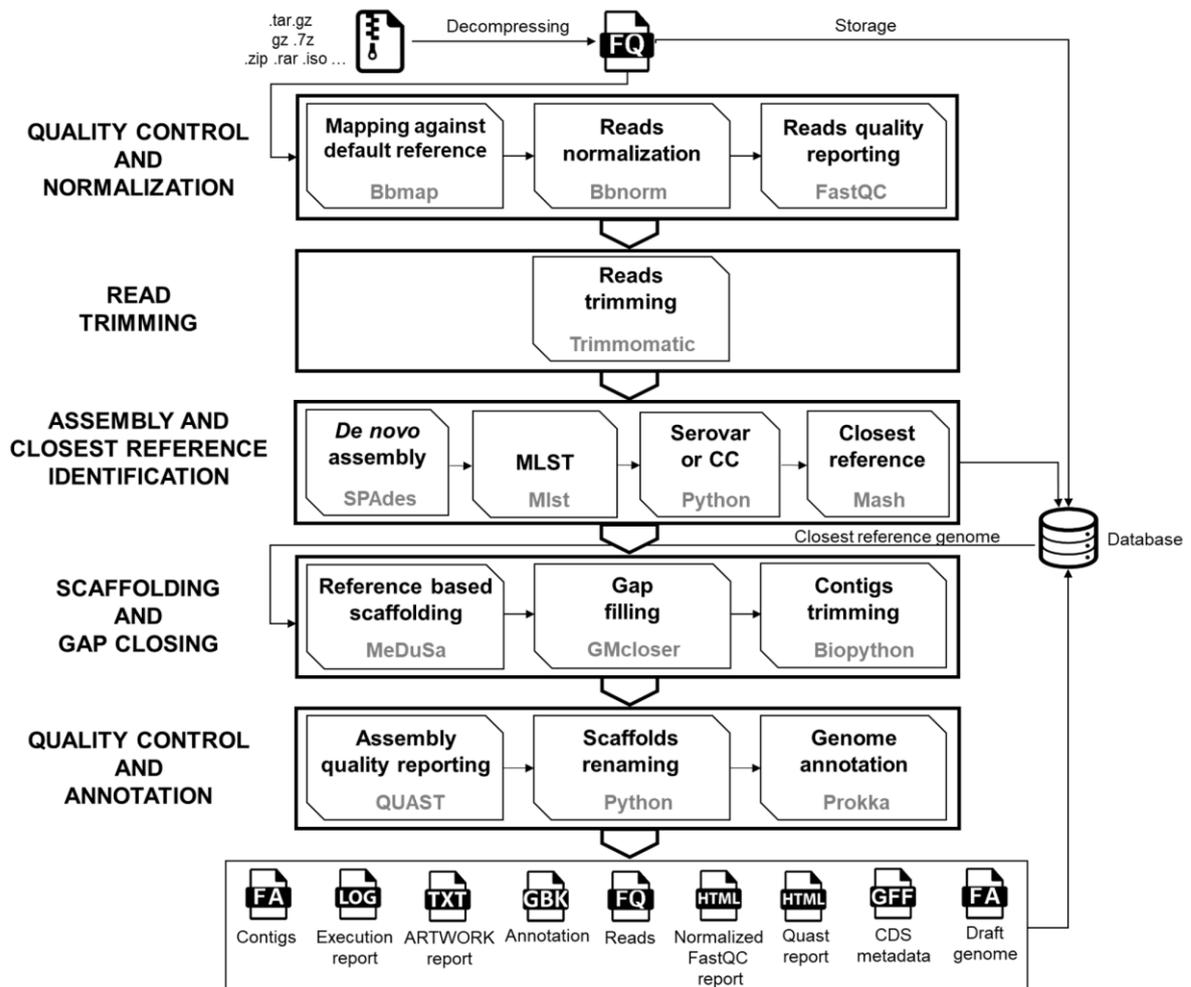


Figure 9 : Outil « assembly of reads and typing workflow » (ARTwork) développé par la mission génome analyse modélisation risque (GAMeR) au sein du laboratoire de sécurité des aliments (LSAL) pour assembler *de novo* des génomes bactériens [96–99, 102, 104, 128]



II.1.b. Intégration des développements à un réseau Linux

Suite à mes travaux de postdoctorat, j'ai intégré la mission GAMeR du LSAL et j'ai amélioré l'algorithmie de cet outil iVARcall2 (Figure 8). Par la suite, un collègue (Arnaud Felten) s'est chargé de l'automatisation et de l'adjonction d'algorithmes d'optimisation [101], afin de transmettre ce dernier à des étudiants et collaborateurs du LSAL. Toujours avec ce même collègue, nous avons aussi développé l'outil ARTwork depuis 2015 qui est aujourd'hui toujours en cours d'amélioration (Figure 9). Depuis 2018, un autre collègue de la mission GAMeR (Ludovic Mallet) a adapté ces deux outils iVARcall2 et ARTwork à un système de parallélisation des calculs (c.à.d. solution open source d'ordonnancement de tâches informatiques « slurm » sur cluster de calcul). Ceci nous a permis de diminuer les temps d'exécutions afférents de ces outils iVARcall2 et ARTwork. Avec une capacité de stockage de 74 TB fournie par la DTI en 2015 et 2019 (utilisée à 78-87% en date du 11 février 2020), nous sommes aujourd'hui en mesure d'appliquer ces outils iVARcall2 et ARTwork en environ 400 minutes pour traiter 96 génomes bactériens. Plus précisément, le cluster de calcul développé par la mission GAMeR est constitué de 7 nœuds, l'un présentant 40 cores et 189 GB de mémoire, et les 6 autres 50 cores et 378 GB de mémoire (Figure 10). Avec le concours significatif de Thomas Texier de la DTI, le réseau Linux développé par la mission GAMeR a dernièrement été synchronisé avec le réseau Windows de l'ANSES via l'application Samba. Ceci permet aux utilisateurs d'accéder aux données génomiques via le réseau Windows supporté par la DTI de l'ANSES ou le réseau Linux développé par la mission GAMeR.

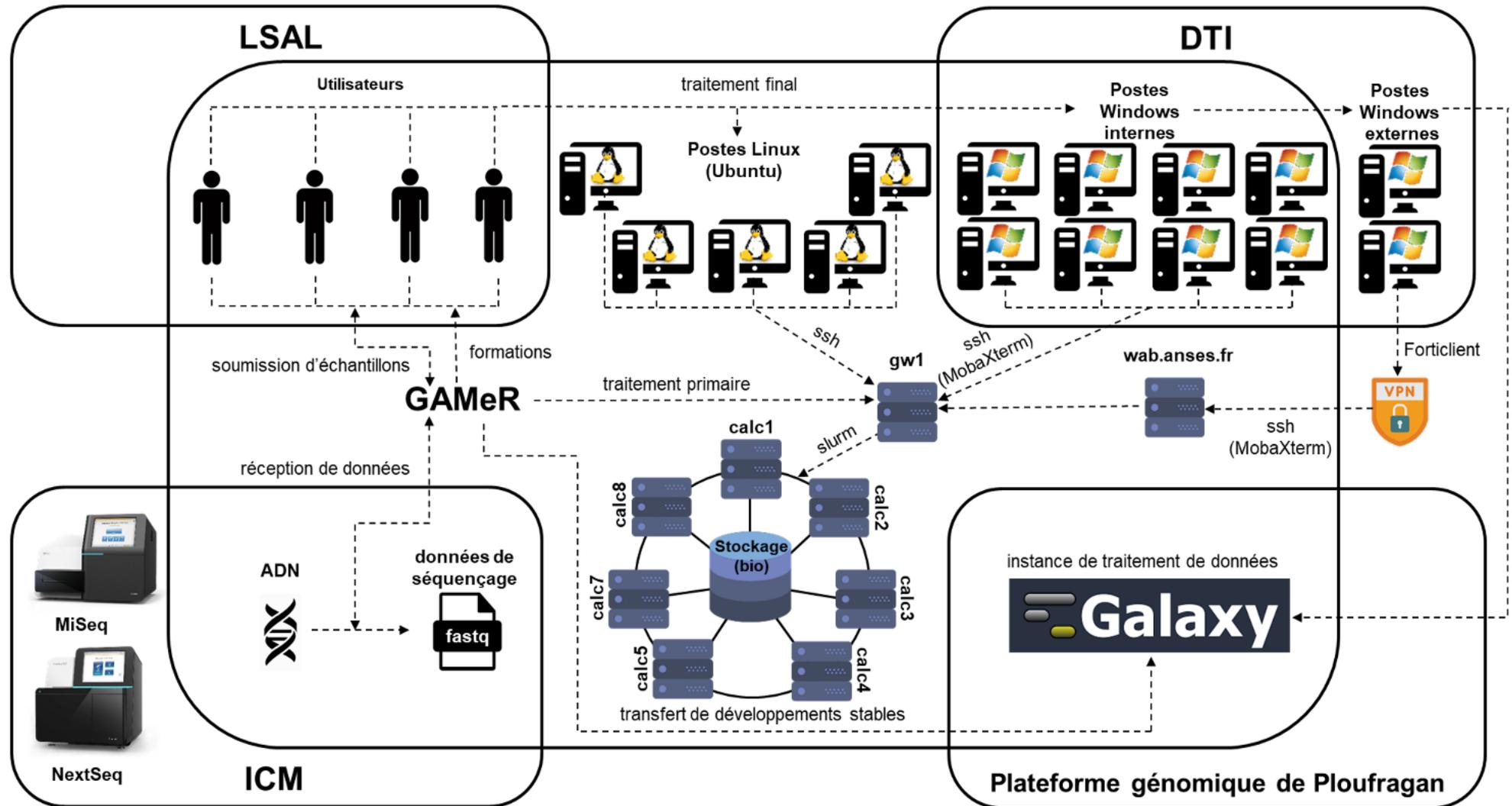


Figure 10 : Réseau Windows-Linux développé par la mission génome analyse modélisation risque (GAMeR) au sein du laboratoire de sécurité des aliments (LSAL) en collaboration avec la direction technique et informatique (DTI), la plateforme génomique de Ploufragan et l'institut du cerveau et de la moelle épinière (ICM)

II.1.c. Transmissions des données et développements stables

Ces deux outils iVARcall2 et ARTwork, ainsi que des promulgations de formations annuelles théorique et pratique à l'utilisation du réseau Linux que la mission GAMeR a construit (Figure 10), ont permis de créer des collaborations avec des étudiants en doctorat du LSAL. Ces étudiants en doctorat ont travaillé sur *Salmonella* (Yann Sévellec) [96, 98, 100], *Listeria* (Lena Fritsch et Clémentine Henri) [94, 99] et *Clostridium* (Abdelrahim Abakabir Mahamat) [97]. Ces transferts de développements stables ont aussi permis d'apporter un appui aux postdoctorants du LSAL travaillant sur *Listeria* (Federica Palma) [128] et *Bacillus* (Mathilde Bonis) [104]. Cette année 2019, c'est une étudiante en Master du Liban (Marie-Noel Mansour) qui est venue utiliser l'organisation du NGS et les outils analytiques mis à disposition par la mission GAMeR. Ses objectifs furent d'étudier l'évolution des « *Salmonella* pathogenic islands » (SPI), plasmides et gènes d'antibiorésistance du sérovar *Salmonella* Enteritidis posant des problèmes sanitaires dans les filières aviaires du Liban (Annexe 1). En plus de ces collaborations avec des doctorants et postdoctorants du LSAL, j'ai aussi développé d'autres projets collaboratifs entre la mission GAMeR et des collègues permanents d'autres équipes du LSAL (Sabrina Cadel-Six et Patrick Fach). Ces collaborations ont donné lieu à une publication sur l'estimation de la justesse de cibles moléculaires à l'échelle génomique (c.à.d. l'outil GTEvaluator) [95] (Annexe 7) et des demandes de brevets concernant l'identification de kmers caractéristiques de sous-ensembles de génomes (c.à.d. l'outil GTFinder) [91] (Annexe 8). Plus précisément, le réseau Linux développé par la mission GAMeR du LSAL a été construit en collaboration avec trois partenaires principaux (Figure 10). En premier lieu, l'institut du cerveau et de la moelle épinière (ICM) qui est notre partenaire en NGS (Yannick Marie). En second lieu, la DTI, le gestionnaire du réseau de l'ANSES (Thomas Texier et Pierre-Yves Letournel). Pour sa part, la plateforme génomique de Ploufragan (Yannick Blanchard) a la responsabilité d'une instance Galaxy de traitement de données pour des utilisateurs nécessitant une interface graphique de développement d'outils analytiques. À l'instar de la plateforme génomique de Ploufragan, la mission GAMeR propose un réseau Linux de traitement de données pour des utilisateurs de cluster de calcul. Via le système d'environnement Conda, la mission GAMeR installe ses développements d'outils dans le cluster de calcul qu'elle gère. Par ailleurs, la mission GAMeR transmet ses développements les plus stables à la plateforme génomique de Ploufragan qui a la charge de les implémenter dans le système Galaxy qu'elle gère. Pour finir, la mission GAMeR, en particulier mon collègue Arnaud Felten, et deux étudiants en cursus d'apprentissage en bioinformatique Kévin Durimel et Pauline Barbet, ont aussi développé une base de donnée recueillant les génomes traités par la mission GAMeR (Figure 11). Ils ont aussi développé une interface graphique d'accès aux données prétraitées (adresse interne à l'ANSES : Adresse web 5) pour les utilisateurs du LSAL. Ces utilisateurs peuvent ensuite poursuivre le traitement de ces données via le cluster de calcul géré par la mission GAMeR (adresse interne à l'ANSES : Adresse web 6 sur Windows ou « ssh -X ansesID@sas-vp-lsgw1 » sur Linux), l'instance Galaxy gérée par la plateforme génomique de Ploufragan ou tout autre système de traitement de données en génomique (Tableau 1).

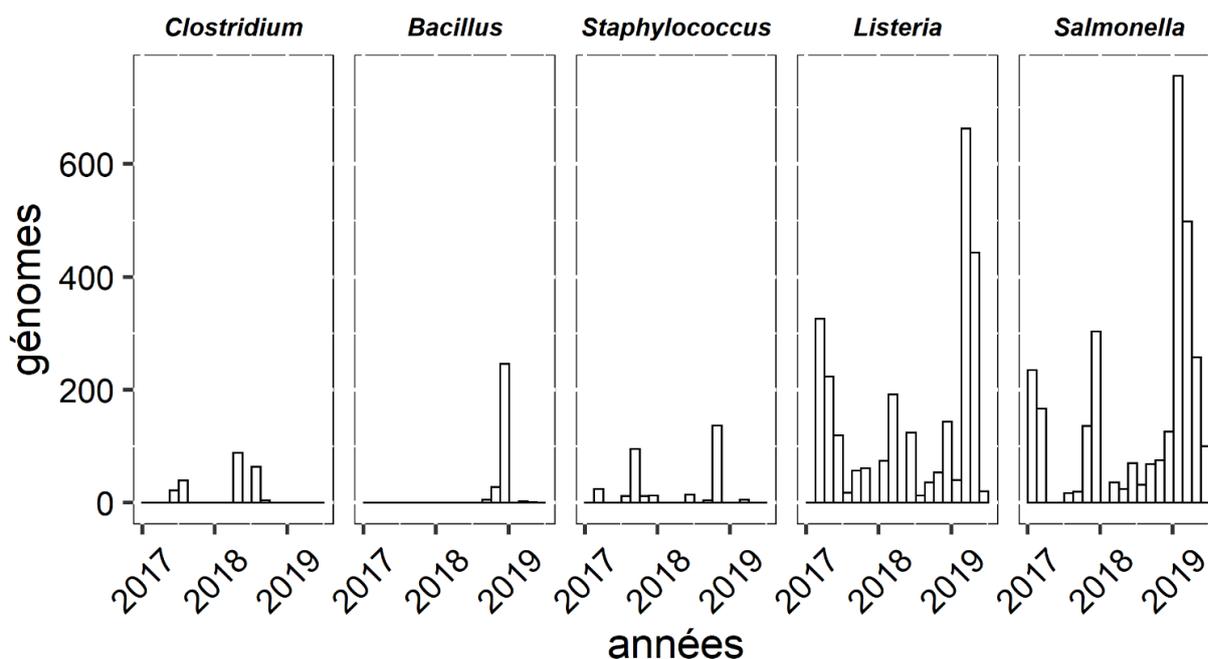


Figure 11 : Évolution du nombre de génomes intégrés à la base de données de la mission génome analyse modélisation risque (GAMeR) provenant des genres bactériens d'archives internationales ou isolés et collectés au laboratoire de sécurité des aliments (LSAL) : *Clostridium* (n = 216), *Bacillus* (n = 283), *Staphylococcus* (n = 316), *Listeria* (n = 2 605) et *Salmonella* (n = 2 919) en date du 10 juillet 2019 (n = 6 339)

II.2. Direction de projets propres

En plus des travaux de collaboration précédemment présentés (chap. II § II.1), j'ai aussi entrepris dans le cadre de la mission GAMeR de concevoir et diriger mes propres projets de recherche (chap. II § II.2.a), ainsi que des activités associées à ces derniers (chap. II § II.2.b).

II.2.a. Travaux de recherche dirigés

Outre mes travaux de doctorat (chap. II § I.2) et postdoctorat (chap. II § I.1) pour lesquels j'ai été en charge d'étudiants en master et doctorat travaillant sur des projets dont j'avais défini les sujets avec ma hiérarchie (Annexe 1), j'ai dernièrement été à l'initiative de plusieurs projets de recherche depuis 2015 au sein de la mission GAMeR du LSAL. Les conceptions, organisations et direction de ces projets de recherche pilotés par la mission GAMeR impliquaient le management de collègues, ainsi que d'étudiants en master, doctorat et postdoctorat (Tableau 3).

Tableau 3 : Outils analytiques développés par la mission génome analyse modélisation risque (GAMeR) entre 2014 et 2019

Outil	Objectif *	Référence
GTEvaluator	Estimation de la sensibilité et spécificité de séquences moléculaires connues vis-à-vis de groupes de génomes	[95]
GTFinder	Identification des combinaisons de séquences moléculaires sensibles et spécifiques de groupes de génomes	[91]
iVARCall2	Revue des SNPs et InDels du coregénomme bactérien par « variant calling analysis »	[96, 97, 99, 101, 102, 104, 118, 120, 128]
FixedVar	Identification des SNPs et InDels du coregénomme bactérien spécifiques des nœuds d'une reconstruction phylogénomique	[101]
ARTwork	Assemblage <i>de novo</i> de génomes bactériens, ordonnancement de contigs, comblement des écarts et annotation de gènes	[96–99, 102, 104, 128]
QuickPhylo	Rapide reconstruction phylogénomique bactérienne par sélection de kmers et estimations des distances de Jaccard	[92]
fastGOES	Association des annotations de mutations bactériennes à des ontologies de gènes et tests d'enrichissement de ces ontologies par GOEA	[101, 102]
microbial-GWAS	Association des SNPs et InDels du coregénomme et gènes du génome accessoire à des traits phénotypiques par GWAS	[102]
Pairwise-FBO	Supports statistiques non-paramétriques pour associer des génomes à des FBO aux échelles SNPs, gènes, kmers et allèles	[92]
COMPASS	Constitution d'une base de données exhaustive et compréhensive de plasmides complètement assemblés	[103]
GAMeRdb	Base de donnée MongoDB de stockage de génomes et variants bactériens	en cours
NauRa	Création automatique de bases de données alléliques des toxines de <i>Staphylococcus</i> pour de la caractérisation de routine	en cours
GENIAL	Création automatique de bases de données de blast et revues des présences et absence de séquences dans une collection de génomes	en cours
Pan-SNP	Inférence phylogénomique des SNPs à l'échelle du pangénomme	en cours

* single nucleotide polymorphisms (SNPs); small insertions/deletions (InDels); genome wide association study (GWAS), gene ontology enrichment analysis (GOEA); foodborne outbreak (FBO)

L'un des projets pour lequel j'ai apporté des contributions majeures de conception et encadrement, concernent notamment le développement d'une méthode d'investigation des épidémies d'origine alimentaire. J'ai réalisé ce projet en lien avec mon responsable hiérarchique (Michel-Yves Mistou) et en collaboration avec le centre national de référence (CNR) des *Escherichia coli*, *Shigella* et *Salmonella* [92]. Ces travaux ont visé à développer une approche statistique d'assignation de génomes à des TIAC (c.à.d. l'outil Pairwise-FBO) sur la base de tests non-paramétriques (Wilcoxon-Mann-Whitney, Kolmogorov-Smirnov, Kruskal-Wallis) comparant les différences par paires de SNPs, gènes, kmers et allèles (cgMLST et wgMLST) [92] (Figure 12 et Annexe 9).

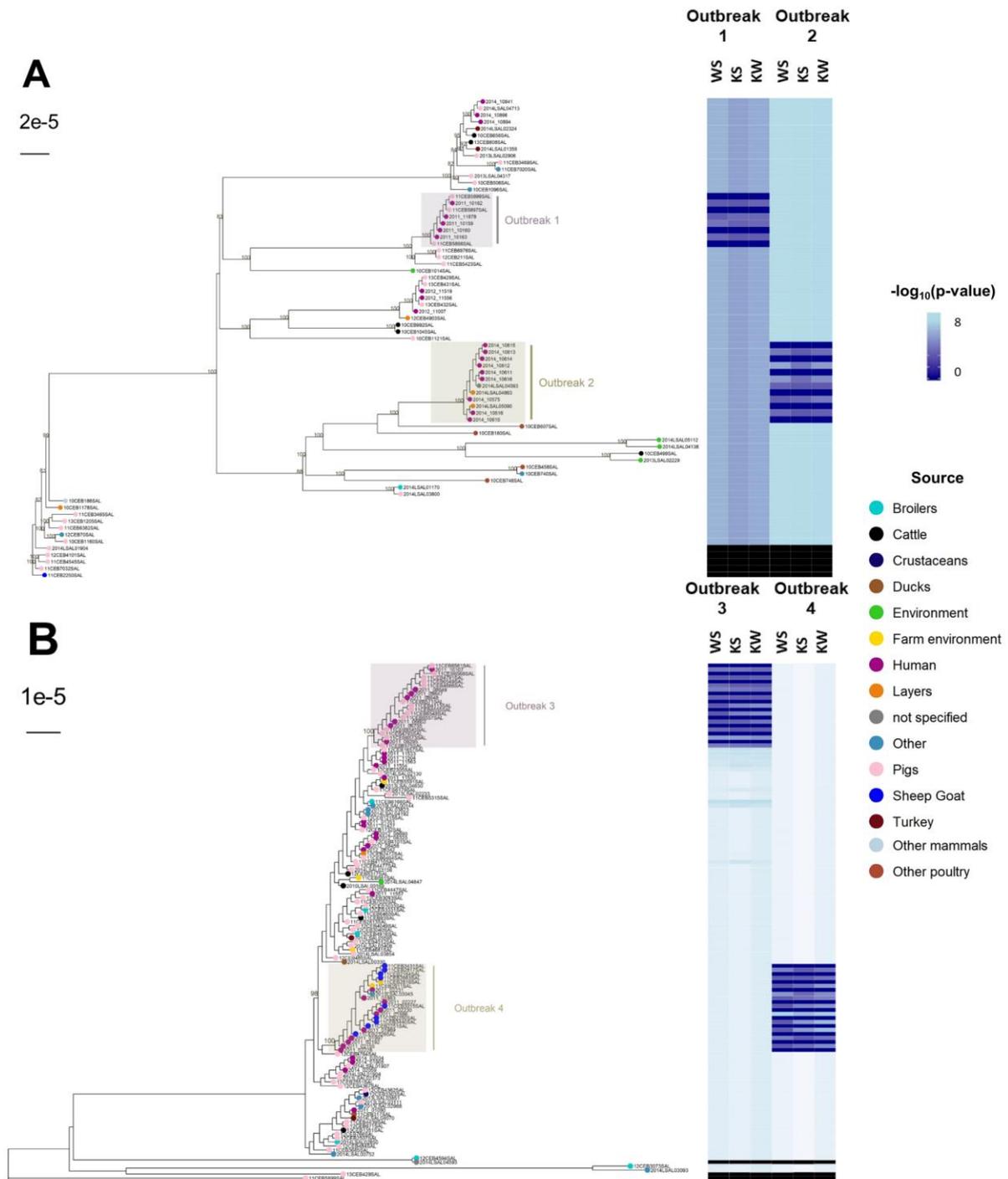


Figure 12 : Inférences phylogénomiques des SNPs du coregénom de souches de *Salmonella enterica* spp. *enterica* (n = 192 génomes) impliquées dans des épidémies des serovars *S. Typhimurium* (A) and *S. 1,4,[5],12:i:-* (B) et résultats des tests non-paramétriques Wilcoxon rank sum (WS), Kolmogorov-Smirnov (KS) and Kruskal-Wallis (KW) [92]

De plus, deux autres projets présentant des contributions majeures de ma part en conception et encadrement, ont visé à développer deux approches de décryptage de l'adaptation de *Salmonella* aux animaux. J'ai réalisé ces deux projets en lien avec Michel-Yves Mistou, Arnaud Felten [101] et une doctorante (Méryl Vila Nova) [102] que j'ai co-encadré en collaboration avec une équipe de l'INRA : l'unité de recherche mathématiques et informatique appliquées du génome à l'environnement (MaIAGE). La première approche consiste à identifier les mutations du coregénome (variants du coregénome homoplastiques ou non-homoplastiques : SNPs et InDels) fixées aux nœuds de l'inférence phylogénomique d'une population clonale ou panmictique (c.à.d. l'outil PhyloFixedVar) [101] (Figure 13 et Annexe 10).

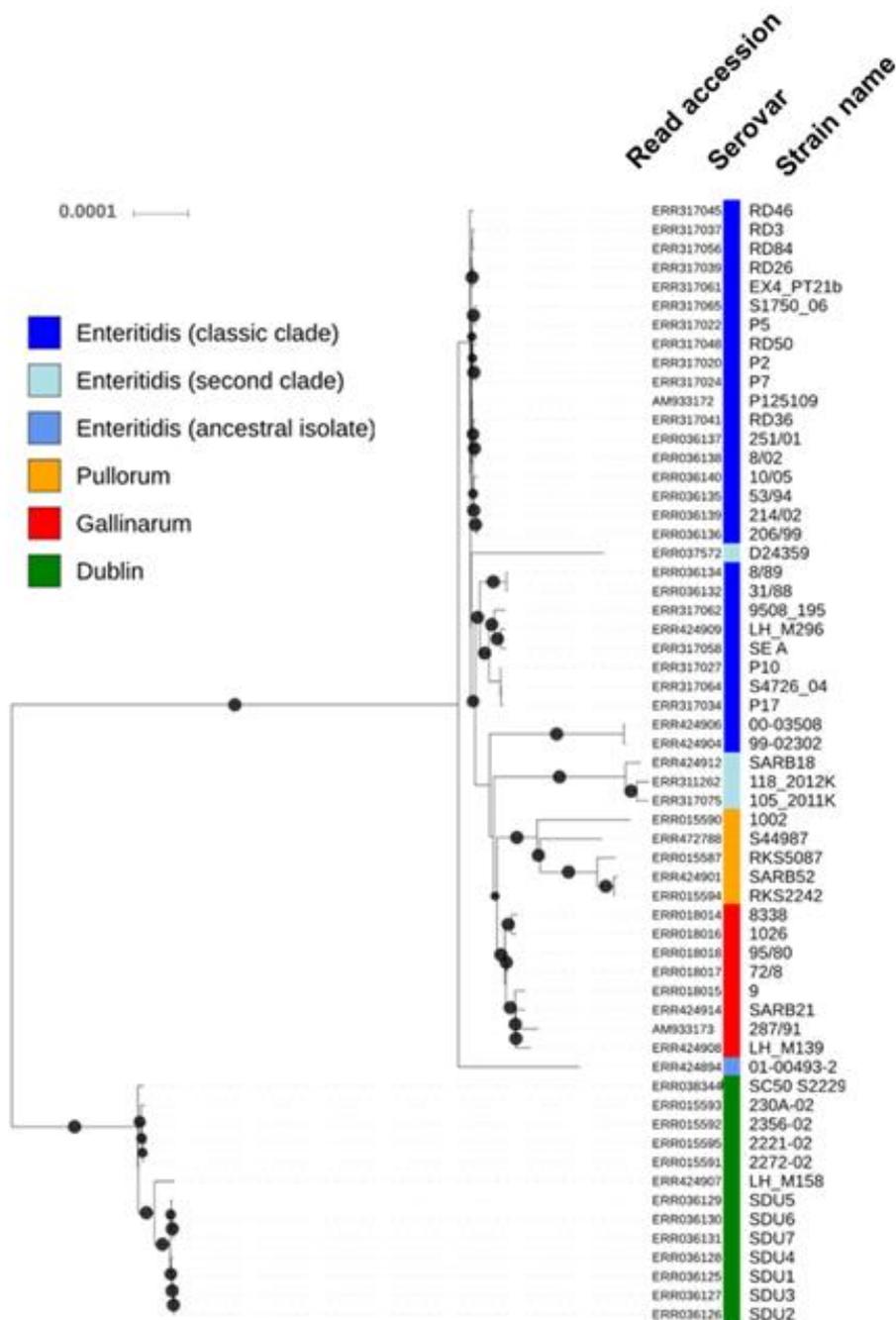


Figure 13 : Inférence phylogénomique des SNPs du coregénome de souches de *Salmonella enterica* spp. *enterica* (n = 73 génomes) des serovars Dublin, Enteritidis, Pullorum et Gallinarum, respectivement des modèles d'adaptation aux bovin, à de multiples hôtes et des hôtes aviaires, visant à identifier les variants fixés aux nœuds des divergences (c.à.d. l'outil PhyloFixedVar) [101]

La seconde approche menée dans le cadre d'une thèse cofinancée ANSES-INRA (Typautobac) vise à identifier les mutations (gènes accessoires et variants du coregénom) associées à des traits phénotypiques binaires dans une population clonal ou panmictique (c.à.d. l'outil microbial-GWAS) par « genome wide association study » (GWAS) [102] (Figure 14 et Annexe 11).

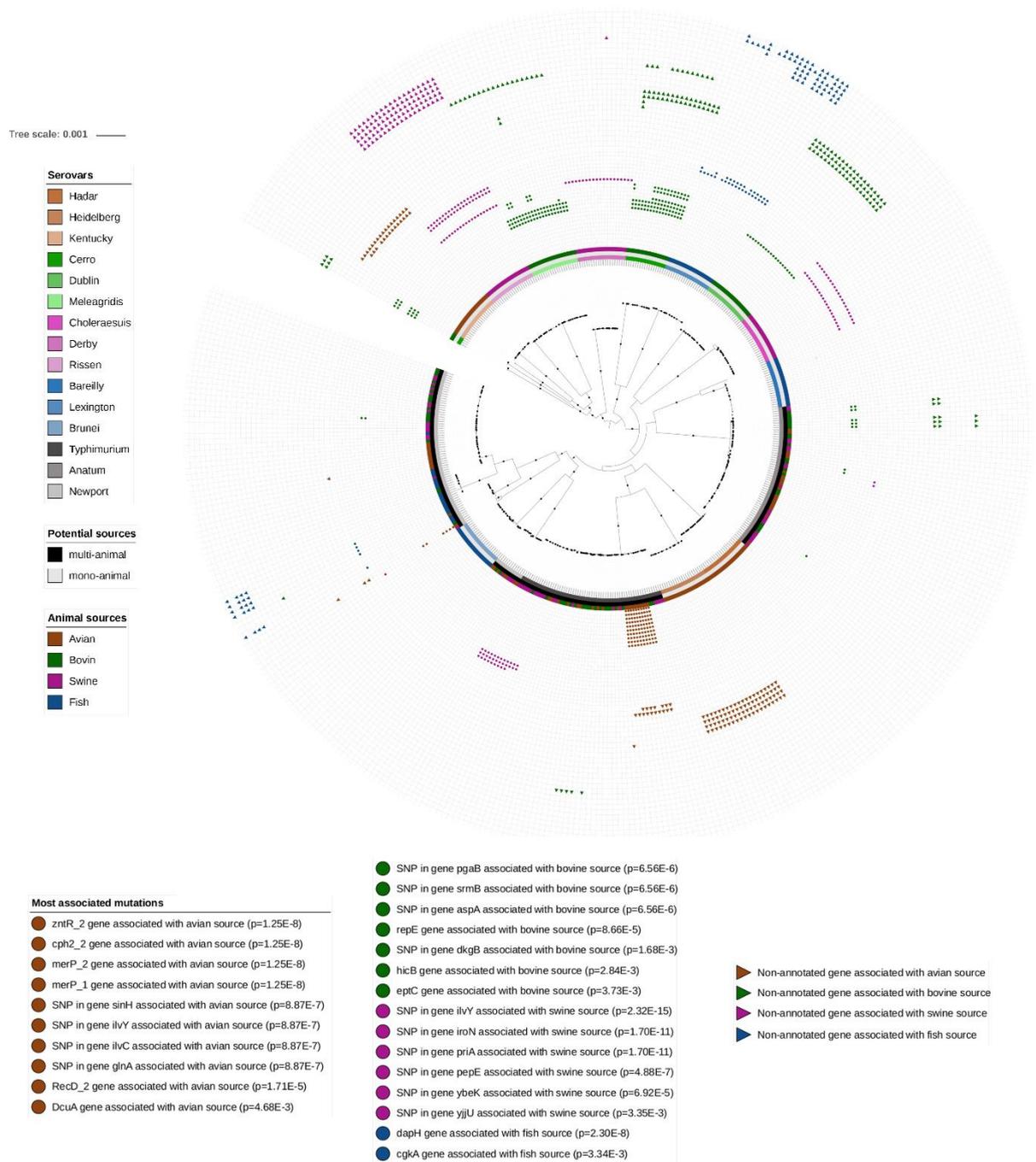


Figure 14 : Inférence phylogénomique des SNPs du coregénom de souches de *Salmonella enterica* spp. *enterica* (n = 440 génomes) de serovars considérés comme adaptés à des sources animales mono- et multi-hôtes et mutations associées par « genome wide association study » (GWAS) (c.à.d. l'outil microbial-GWAS) [102]

Ces deux approches d'identification de mutations d'intérêt ont aussi été couplées à un autre outil développé par la mission GAMeR (c.à.d. l'outil fastGOES), afin d'identification les voies métaboliques enrichies correspondantes par « gene ontology enrichment analysis » (GOEA) [101, 102] (Figure 15 et Annexe 12). Sous ma responsabilité et dans le cadre de sa thèse de doctorat, Méryl Vila Nova a encadré à ce sujet un étudiant de master (Kévin La).

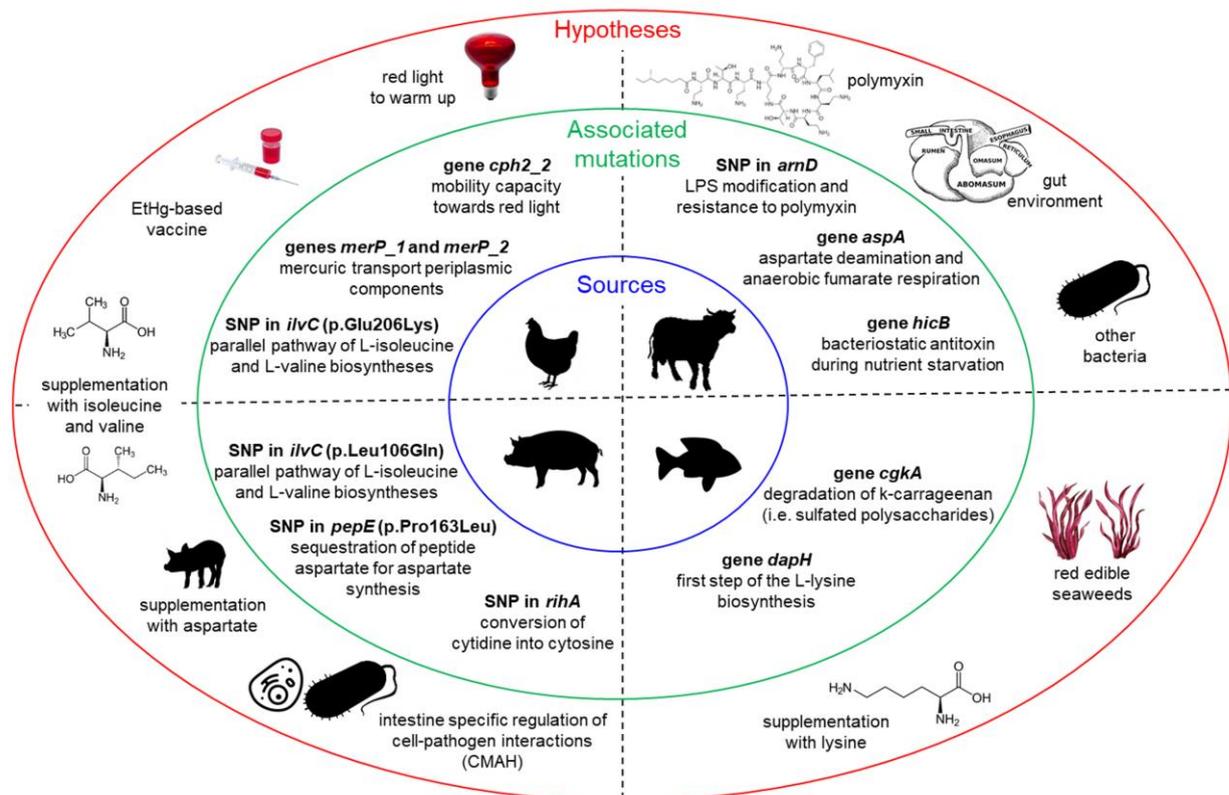


Figure 15 : Mutations du génome accessoire et du coregénomme associées par « genome wide association study » (GWAS) (c.à.d. l'outil microbial-GWAS) à des sources animales mono- et multi-hôtes de sérovars de *Salmonella enterica* spp. *enterica* (n = 440 génomes) et voies métaboliques correspondantes identifiées par « gene ontology enrichment analysis » (GOEA) (c.à.d. l'outil fastGOES) [102]

Dans le cadre du projet EJP RaDAR, le postdoctorant Pierre-Emmanuel Douarre, avec l'aide de Ludovic Mallet, mène des travaux sur les plasmides bactériens qui cimentent actuellement la cohésion des membres de la mission GAMeR. Une base de données de plasmides (Figure 16) a été construite et nommée « comprehensive and complete plasmid database » (COMPASS) et nous permet de comprendre les relations existantes entre mobilisations plasmidiques, Phyla bactériens et résistances aux antibiotiques [103].

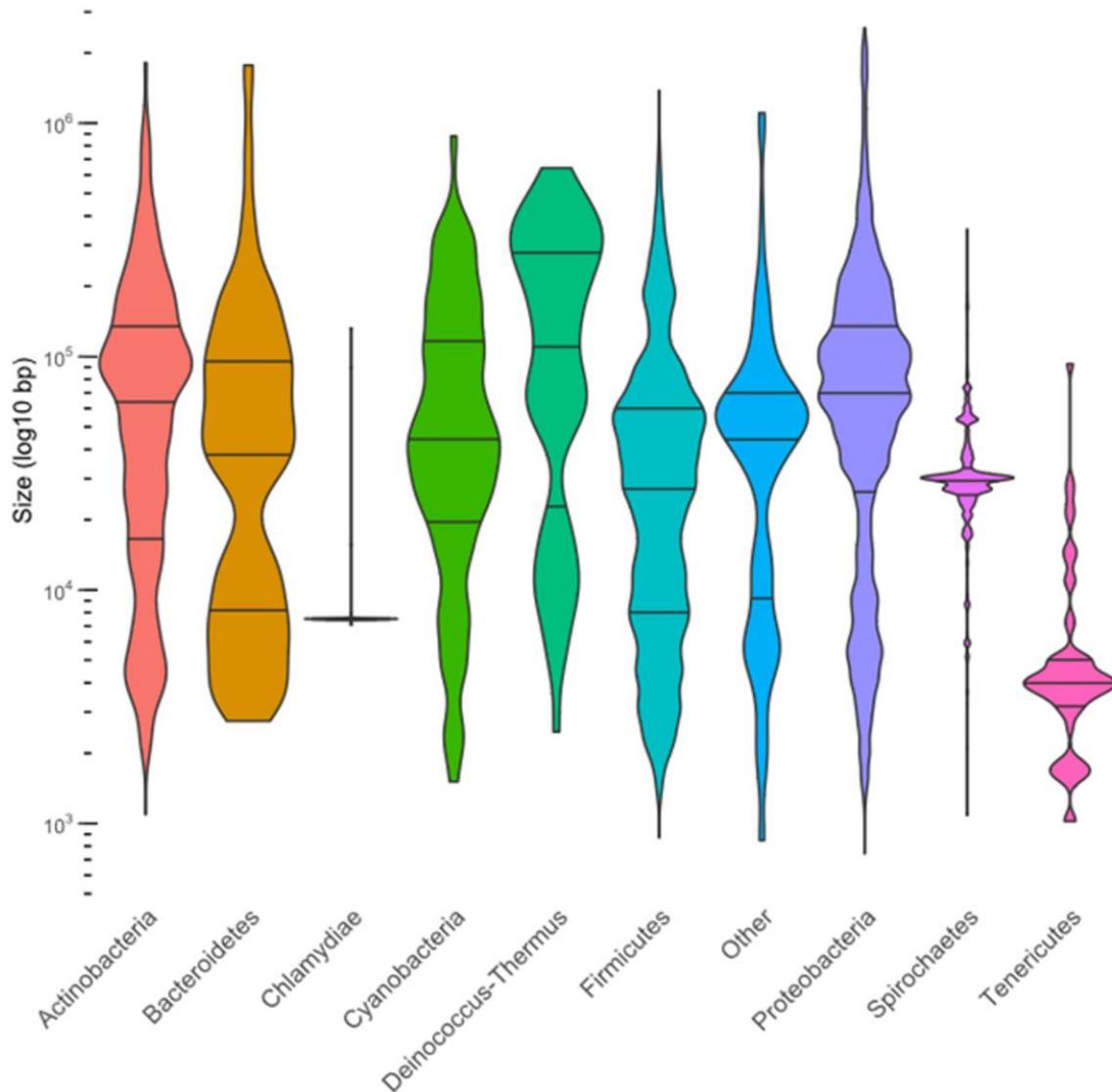


Figure 16 : Distribution des tailles de plasmides (log₁₀ de paires de bases) en fonction des Phyla bactériens de la base de plasmides (n = 12 084) « comprehensive and complete plasmid database » (COMPASS) [103]

II.2.b. Supports d'activités

Certains supports d'activités ont été nécessaires à la bonne conduite des projets de recherche que j'ai portés (chap. II § II.2) et ont impacté positivement le déroulement de projets collaboratifs (chap. II § I). J'ai donc aussi depuis 2015 pris la charge au sein de la mission GAMeR, d'activités pour implémenter de la génomique au LSAL. Ces supports d'activités furent des créations de marchés publics (chap. II § II.2.b.i), des négociations d'investissement (chap. II § II.2.b.ii), des animations de la recherche (chap. II § II.2.b.iii) et le maintien de réseaux d'étudiants (chap. II § II.2.b.iv).

II.2.b.i. Marchés publics et procédures afférentes

Dans l'objectif d'obtention de données NGS de hautes qualités, j'ai tissé des liens avec plusieurs centres dans le domaine afin de comparer leurs prestations analytiques (GATC Biotech, Genewiz, Genoscreen, ICM, Integragen, Microsynth, Eurofins, Intragen, Genomicos). Cette connaissance du marché m'a conduit en 2016 puis en 2018 à créer et ouvrir deux marchés à procédure adaptée (MAPA) pour le séquençage haut débit complet de génomes bactériens en collaboration avec la DAF de l'ANSES. Ces deux MAPA (n°16A000026 et n°XBDC000257-18FCS024) ont permis aux équipes du LSAL d'externaliser le séquençage complet de génomes. Le dernier MAPA en date est valable pour 2 ans et renouvelable deux fois pour une période totale de 6 ans (Tableau 4). En parallèle de la construction du réseau Linux au sein de la mission GAMeR, je me suis aussi chargé de l'harmonisation des procédures du laboratoire concernant l'extraction d'ADN (procédure Wizard de Promega et automate de Kingfisher) et la soumission des échantillons (Annexe 13). Ces soumissions d'échantillons ont été gérées par la mission GAMeR entre 2015 et 2018 concernant les communications et transferts de données avec notre partenaire de séquençage (ICM). En 2018, je fus en charge du transfert de l'activité de gestion des soumissions d'échantillons à séquencer et prétraiter vers des correspondants NGS des unités *Salmonella* et *Listeria* (SEL) (Maroua Sayeb et Emeline Cherchame) et *Staphylococcus*, *Bacillus*, *Clostridium* (SBCL) (Déborah Merda). De 2015 à 2018 et toujours actuellement en 2019, la mission GAMeR se charge de l'exécution systématique des outils iVARcall2 et ARTwork (chap. II § II.1.a, § II.1.b et § II.1.c).

Nombre d'échantillons d'ADN génomique	Cadre analytique	Délais de séquençage et remise de résultats (jours ouvrables)	Prix remisé (€ hors taxe)	Prix remisé unitaire (€ hors taxe)
192	recherche	15	12 090	62,9
96	recherche	15	8 161	85,0
96	recherche	10 à 15	8 161	85,0
5 à 32	urgence sanitaire	10	2 230 à 3 310	446 à 103
5 à 32	urgence sanitaire	5 à 10	2 230 à 3 310	446 à 103
2 à 4	urgence sanitaire	10	1 530 à 1 610	765 à 402
2 à 4	urgence sanitaire	5 à 10	1 530 à 1 610	765 à 402

Tableau 4 : Prestations du marché à procédure adaptée (MAPA) portant sur le séquençage haut débit (Librairies Nextera XT et séquenceur Illumina 2 x 150 bases Nextseq500) de génomes bactériennes (taille de 2-7 Mb et couverture minimale théorique de 30X), conclu en 2018 (n°XBDC000257-18FCS024) entre l'agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (ANSES) et l'institut du cerveau et de la moelle épinière (ICM), valable pour 2 ans, et renouvelable deux fois pour une période totale de 6 ans

II.2.b.ii. Mise en place des moyens matériels

Comme observé par d'autres laboratoires, les capacités de stockage et ressources de calcul sont actuellement les goulots d'étranglement de l'implémentation de la génomique (intro. § II.2). À mon arrivée au LSAL en 2015, trois ordinateurs Linux à la disposition de la mission GAMeR ont permis d'initier notre activité d'intégration de la génomique. Des réunions et négociations semestrielles avec la DTI, ont permis de financer deux ordinateurs supplémentaires pour supporter les travaux des étudiants de la mission GAMeR et les formations à l'utilisation du réseau Linux que nous promulguons au sein du LSAL. En 2016 et 2017, la DTI a mis en place une zone de stockage et un serveur de calcul pour la mission GAMeR additionnés de six autres fournis par l'INRA de Toulouse en 2019. Aujourd'hui, cette zone de stockage est liée au cluster de calcul et le réseau Linux de la mission GAMeR est synchronisé au réseau Windows de l'ANSES (Figure 10). Ceci permet aux utilisateurs du cluster de calcul proposé par la mission GAMeR de s'y connecter via le réseau de l'ANSES avec un terminal d'ordinateur Linux ou un émulateur de terminal Linux sur un ordinateur Windows (MobaXterm). De l'extérieur à l'ANSES, les utilisateurs du cluster de calcul proposé par la mission GAMeR peuvent aussi s'y connecter via un accès « virtual private network » (VNP) (Forticlient) sous réserve qu'ils aient un compte informatique ANSES.

II.2.b.iii. Animation de la recherche

Entre 2015 et 2019 et en collaboration avec l'UZB du LSAN de Maisons-Alfort, j'ai contribué au sein de la mission GAMeR (Figure 17) à l'organisation des séminaires trimestriels entre les unités des deux laboratoires, ainsi que des séminaires annuels avec des invités extérieurs, dans le cadre du DEBUG (chap. I § III.2.d). Tout en organisant les calendriers d'intervention, mes activités étaient d'identifier et inviter les participants au DEBUG. Dans un autre registre et en parallèle de la construction du système de traitement de données (chap. II § II.1.a et § II.1.c), j'ai piloté la participation de la mission GAMeR à trois essais inter-laboratoires en génomique bactérienne organisés par le « global microbial identifier » (GMI). J'ai par la suite transféré cette activité aux correspondants NGS des unités SEL et SBCL (chap. II § II.2.b.i). À la fois en tant que participant, intervenant et présentateur invité, j'ai aussi participé et continue à participer à des congrès nationaux et internationaux (Annexe 1). Ma participation à l'animation de la recherche s'est aussi exprimée par 41 révisions ou relectures de manuscrits soumis à éditeurs de rang A (Annexe 1).

II.2.b.iv. Participation à la formation par la recherche

De par ma connaissance des établissements de formations entre 2008 et 2015, j'ai enseigné la microbiologie à l'institut universitaire et technologique (IUT) de Créteil et l'université Paris Diderot durant mon monitorat d'enseignement en parallèle de mon doctorat. J'ai aussi enseigné la bioinformatique sous forme de vacances d'enseignement à ParisTech, pour un total de 177 heures d'équivalents travaux dirigées (Annexe 1). Concernant ces activités d'enseignement, j'ai obtenu en France des qualifications aux fonctions de maître de conférences des universités en section 65 biologie cellulaire, 67 biologie des populations et écologie, ainsi que 68 biologie des organismes (Annexe 1). Au Canada, j'ai obtenu un certificat de sélection du Québec en tant qu'assistant d'enseignement et recherche postsecondaire (Annexe 1). Contribuant aussi aux relations de confiance avec des étudiants en cursus de doctorat entre 2016 et 2019, je me suis chargé de l'organisation des deux réunions de comités de thèse de l'étudiante Méryl Vila Nova. J'ai aussi eu l'opportunité de faire partie de deux jurys de thèse des doctorants Franck Cerutti et Damien Richard, ainsi que des deux réunions de comités de thèse de la doctorante Maëllys Kevin (Annexe 1). Mes appartenances à la société française de bioinformatique (SFBI), la société chimique américaine (ACS), la société française de microbiologie (SFM), la société française de mycobactériologie (MycoClub) et l'association française d'écologie microbienne (AFEM), sont aussi des éléments me

permettant de conserver des réseaux d'étudiants et de partenaires désireux de travailler en génomique bactérienne (Annexe 1).



Figure 17 : Membres de la mission génome analyse modélisation risque (GAMeR) en juillet 2018 (de gauche à droite et de haut en bas : Kévin La, Julien Pichon, Méryl Vila Nova, Arnaud Felten, Kévin Durimel, Émilie Pasteur, Pauline Barbet, Pierre-Emmanuel Douarre, Ludovic Mallet, Michel-Yves Mistou, Nicolas Radomski)

« Plus nous avançons dans la connaissance et la maîtrise des technologies de la vie, plus nous nous condamnons à être responsables »

Olivier Arnaud

CHAPITRE III : PROGRAMME DE RECHERCHE

Au regard de mon environnement de travail (CHAPITRE I : LIEU D'ACCUEIL) et ma carrière personnelle (CHAPITRE II : BILAN DE CARRIÈRE), le présent chapitre de mémoire est présenté en vue de l'obtention de l'HDR (CHAPITRE III : PROGRAMME DE RECHERCHE). Ce chapitre commence par décrire les adéquations entre mon environnement de travail et ma carrière personnelle. Tous les deux sont orientés vers une numérisation des données de santé publique (chap. III § I.1) et une amélioration continue des pratiques analytiques dans les domaines de la microbiologie (chap. III § I.2). Ensuite, ce chapitre finit par les présentations des projets de recherche que j'envisage conduire pour les années à venir (chap. III § II).

I. Adéquations entre environnement de travail et carrière personnelle

Les concordances de trajectoires entre mon lieu d'accueil et mon parcours de chercheur sont de deux types. D'une part, l'implémentation des travaux de génomique réalisés au cours de ma carrière est de nos jours pleinement prise en compte par les acteurs de santé publique (chap. III § I.1). D'autre part, l'amélioration continue des pratiques analytiques en microbiologie a conduit ma carrière et est une réalité contemporaine en santé publique (chap. III § I.2).

I.1. Implémentation de la génomique par les acteurs de la santé publique

Les recommandations et activités des instances internationales et nationales de différents domaines interdépendants de santé publique, ainsi que les multiples consortia de recherche dans ce secteur, encouragent l'implémentation de la génomique (chap. III § I.1.a). Pour ces raisons, les environnements de travail de l'ANSES (chap. III § I.1.a), du LSAL (chap. III § I.1.b) et de la mission GAMeR (chap. III § I.1.c), tendent vers une implémentation de la génomique pour la santé publique.

I.1.a. Au sein des autorités de santé publique

L'approche multisectorielle « *One Health* » a été définie en 2000 par l'organisation mondiale pour la santé (OMS), puis reprise par l'organisation des nations unies pour l'alimentation et l'agriculture (FAO) et l'organisation mondiale de la santé animale (OIE). Cette approche « *One Health* » est en complète adéquation avec les champs de compétences de l'ANSES. En pratique, cette approche multisectorielle se retrouve dans les activités de l'ANSES en surveillance, référence et recherche appliquées à différentes thématiques de santé publique (chap. I § I.1.c). Effectivement, le gouvernement français, par la voie de sa direction générale de la mondialisation, du développement et des partenariats, soutiennent officiellement cette approche depuis 2011 [129]. Ce soutien se traduit via notamment la mise en place de l'accès au réseau internet à très haut débit sur tout le territoire national d'ici 2022, récemment portée par le projet de la loi Elan sur l'évolution du logement, de l'aménagement et du numérique [130]. Suite au commissionnement du premier ministre en 2015 dans le cadre du plan France médecine génomique 2025, d'importants moyens (670 millions € les 5 premières années) ont été alloués depuis 2016 [71]. Les acteurs du plan France médecine génomique 2025 sont l'alliance nationale pour les sciences de la vie et de la santé (AVIESAN) et le centre de référence, d'innovation, d'expertise et de transfert (CReFIX). Ils sont en charge du développement des procédures, outils et technologies, ainsi que de l'implémentation et la promulgation de stage d'utilisation [71]. Dans le cadre du G7, l'ANSES et ses homologues allemand, danois et coréen, respectivement le « *bundesinstitut für risikobewertung* » (BfR), le « *national food institute, technical university of Denmark* » (DUT Food) et le « *national institute of food and drug safety evaluation* » (NIFDS), ont renouvelé en 2019 leurs engagements à intégrer la génomique au service de la sécurité sanitaire des aliments [131].

1.1.b. Au sein du LSAL de l'ANSES

Les activités actuelles de surveillance, référence et recherche de l'ANSES (chap. I § I.1.c) sont reprises en ces mêmes termes par mon laboratoire d'accueil : le LSAL (chap. I § II.1). Initiée par des préoccupations de santé humaine [132], l'approche « *One Health* » a effectivement été conceptualisée par la suite dans les domaines vétérinaire [133], environnementale [134], puis alimentaire [135]. À ce propos, les modifications d'organisations des laboratoires de microbiologie passant d'une segmentation par pathogènes à une segmentation par méthodes analytiques (intro. § II.4), vont de pair avec l'avènement des technologies NGS. Effectivement, ces profondes modifications d'organisations amènent ces laboratoires à devoir traiter différents aspects des données omiques, telles que par exemple la génomique, la transcriptomique, la métagénomique, la protéomique, la lipidomique, l'épigénétique et bien d'autres encore. Par la création en 2015 de la mission GAMeR, le LSAL s'est doté d'une mission transversale pour intégrer la génomique bactérienne dans ses équipes (Annexe 6).

1.1.c. Au sein de la mission GAMeR

Tous comme les autres missions et équipes du LSAL (Annexe 5), la mission GAMeR est amenée à travailler sur différentes bactéries pathogènes du secteur alimentaire en raison de sa transversalité (chap. II § I.4). L'implémentation de l'épidémiologie à l'échelle génomique en temps-réel par les équipes traditionnelles de contrôle des infections présente de nombreux facteurs limitants. Ces facteurs limitants sont aujourd'hui davantage liés au support du système bioinformatique (intro. § II.2) et à la gestion d'une quantité extrêmement importante de données produites (intro. § II.3), et non plus au coût de séquençage qui diminue continuellement (intro. § II.1) [136]. La mission GAMeR se doit donc de poursuivre ses développements d'un réseau Linux (Figure 10) et d'outils analytiques (Tableau 3) en s'inspirant des plateformes nationale et internationale récemment déployées (Tableau 1).

1.2. Amélioration des pratiques analytiques et intégration des domaines d'étude de la microbiologie

Ma carrière personnelle est totalement en adéquation avec l'approche « *One Health* » soutenue par des organismes internationaux et nationaux (chap. III § I.1.a), par l'implémentation de la génomique à l'ANSES (chap. III § I.1.a) et au LSAL (chap. III § I.1.b) à travers la mission GAMeR (chap. III § I.1.c). Effectivement, mon parcours professionnel est orienté vers l'amélioration des pratiques analytiques (chap. III § I.2.a) en microbiologie humaine, animale, environnementale et alimentaire (chap. III § I.2.b) (Figure 18).

1.2.a. L'amélioration des pratiques analytiques en microbiologie

Mon parcours professionnel vise à améliorer les pratiques analytiques en microbiologie dans l'objectif de les faire diffuser à la communauté scientifique (Figure 18). Effectivement, mes premiers développements ont été réalisés en bactériologie [126], puis en biologie moléculaire de quantification [123–125], de caractérisation phylogénétique [116, 127] et liée au NGS [119]. Par la suite, mes développements se sont focalisés sur la génomique bactérienne incluant la définition de séquences de caractérisation [91, 95, 122, 122], la reconstruction phylogénomique [96, 97, 99, 101, 102, 104, 118, 120, 128] et l'assemblage *de novo* de génomes [96–100, 102, 104, 128]. Plus récemment et dans le cadre de la mission GAMeR, nous avons aussi développé d'autres outils analytiques visant à identifier les variants du coregénomme expliquant les divergences entre lignées [101], détecter les voies métaboliques principalement impactées par des mutations [101, 102] et lister les mutations associées à des traits phénotypiques [102]. D'autres développements nous permettent aussi de reconstruire rapidement des histoires évolutives par sélection de kmers et estimation des distances de Jaccard [92], ainsi que supporter statistiquement l'assignation de génomes à des événements de TIAC [92]. Nos

développements les plus récents nous ont amenés à créer une base de données exhaustive de 12 084 plasmides complètement assemblés [103], ainsi que produire d'autres outils analytiques en cours de développement (Tableau 3). Fervent défenseur des outils bioinformatiques en libre accès, je prévois de continuer à développer de nouvelles méthodes analytiques en génomique bactérienne (chap. III § II.2).

1.2.b. Intégration des domaines d'étude de la microbiologie

En plus de la diversité de mes développements analytiques en bactériologie, biologie moléculaire et génomique (chap. III § II.1.a), j'ai entrepris et financé mes travaux de recherche dans le cadre de différents domaines d'étude de la microbiologie. Plus précisément, j'ai évolué en microbiologie humaine [118–121, 123], animale [127] et environnementale [115, 116, 122, 124–126] concernant l'étude du genre *Mycobacterium*. J'ai de plus aussi évolué en microbiologie alimentaire en travaillant sur les genres *Salmonella* [91, 92, 95, 96, 98, 100–102], *Listeria* [94, 99, 117, 128], *Clostridium* [97] et *Bacillus* [104]. De plus, j'ai dernièrement initié des travaux sur les plasmides au sens large des taxa du règne *Bacteria* [103] (chap. II § I) (Figure 18). Initialement en lien avec des opportunités de financement, ma curiosité pour les différents domaines des sciences en général et de la microbiologie en particulier, explique mes changements de domaines d'étude dans différents laboratoires de microbiologie en France, Etats-Unis et Canada. Ces expériences sont *in fine* en totale adéquation avec l'approche « One Health » visant à avoir un regard global sur les problématiques de santé publique. Ce parcours international m'a amené à apporter des solutions d'un point de vue numérique concernant la gestion des données massives en génomique (chap. III § I.1).

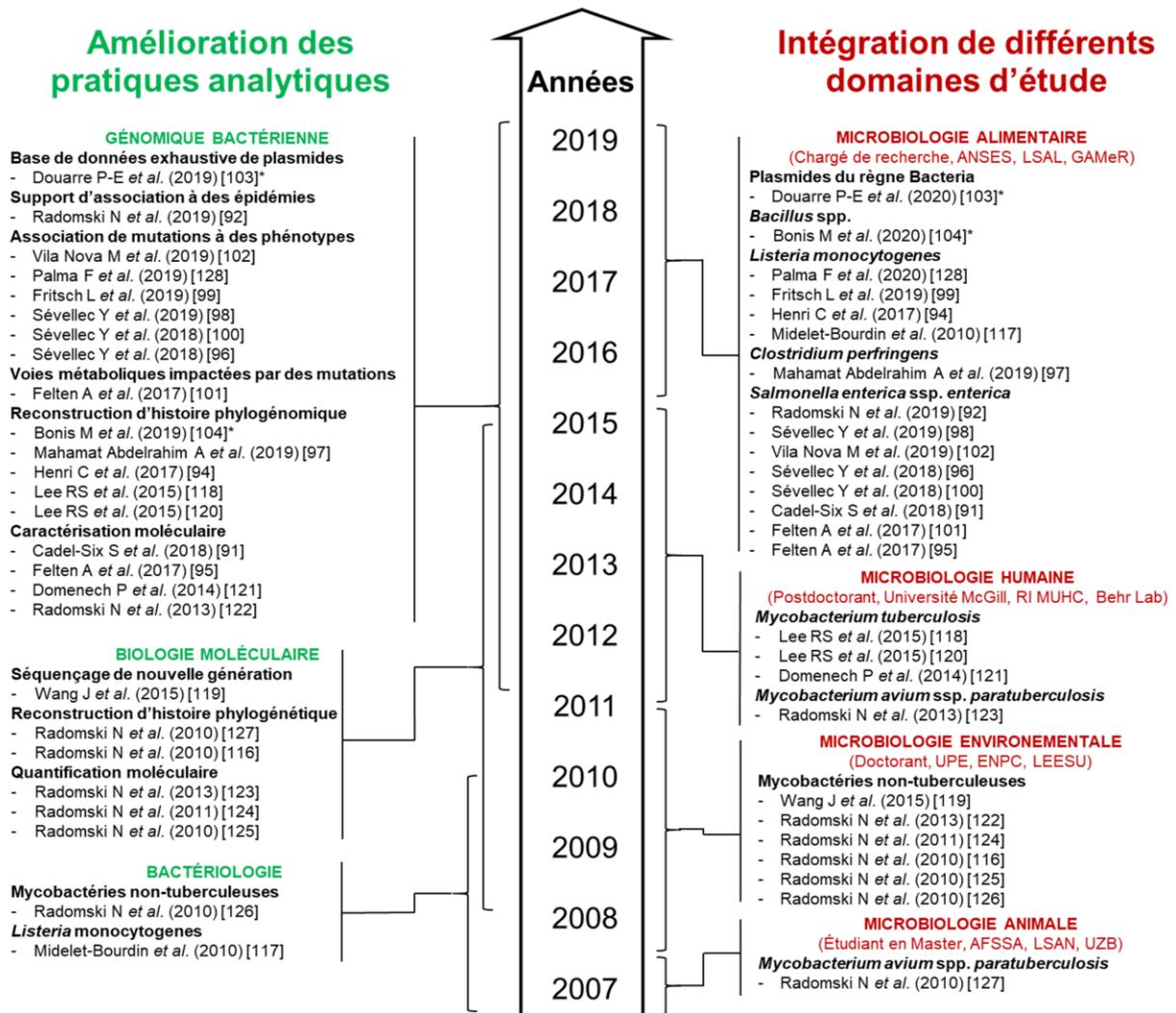


Figure 18 : Articles publiés pendant mon parcours professionnel international orienté vers l'amélioration des outils analytiques en bactériologie, biologie moléculaire et génomique bactérienne dans les domaines de la microbiologie humaine, alimentaire, environnementale et alimentaire (le signe * signifie que le manuscrit est en cours d'évaluation par les pairs)

II. Projets de recherche

Le LSAL au sein duquel je travaille actuellement réalise des activités en surveillance (chap. I § II.1.a), référence (chap. I § II.1.b) et recherche (chap. I II.1.c), dans un objectif de participer à l'approche « *One Health* » (chap. III § I.1.a et I.1.b). En conséquence, je compte développer à l'avenir trois projets dans ces domaines. Le premier projet de surveillance porterait sur l'accessibilité aux membres du LSAL des métadonnées des souches en collection au LSAL (chap. III § II.1). Le second projet utiliserait la génomique pour répondre à des questionnements relatifs au sérovar *Salmonella* Enteritidis dans la filière aviaire (chap. III § II.2). Le troisième projet de recherche en génomique serait appliqué à l'hypothétique origine hydrique des TIAC à *Salmonella* en Europe ou en France (chap. III § II.3).

II.1. **Projet de surveillance : Accessibilité des métadonnées des souches en collection au LSAL dans le cadre de l'intégration à la base de données génomique développée par la mission GAMeR (Projet AcSaMe)**

Sous le nom de « accessibility of sampling metadata » (AcSaMe), le projet de surveillance AcSaMe vise à intégrer les métadonnées d'échantillons soumis pour assemblage et identification des variants dans une base de données pilotée par la mission GAMeR du LSAL.

II.1.a. *Contexte scientifique*

Depuis 2019, l'EFSA propose une interface de programmation d'applications (API) pour que les développeurs puissent travailler sur les données de sécurité alimentaire accessibles au niveau européen (Adresse web 7). L'implémentation de la base de données de génomes développée par la mission GAMeR, surnommée GAMeRdb (Tableau 3 et « stockage bio » de la Figure 10), n'est actuellement reliée à aucune métadonnée concernant les échantillons soumis par les équipes du LSAL (Annexe 13). Ceci ralentit considérablement l'application des outils développés par la mission GAMeR (Tableau 3). Fluidifier le système présent reviendrait soit à transférer les données génomiques dans les bases de données accueillant les métadonnées d'échantillons, soit à transférer ces métadonnées dans la base de données de génome développée par la mission GAMeR (Figure 19). Néanmoins, les bases de métadonnées du LSAL se divisent en plusieurs unités sans harmonisation des formats d'enregistrement (langage R, tableur Excel, base Access, enregistrement papier). À contrario, la base de données génomiques développée par la mission GAMeR présente l'avantage d'être unifiée de par sa conception et directement liée à un cluster de calcul (Figure 10). Tout comme pour les données génomiques, cette intégration des métadonnées d'échantillons dans la base de données développée par la mission GAMeR sous le format MongoDB, permettrait de réaliser rapidement des requêtes en langage Bash ou Python. Ces requêtes envisageables sur tous les projets de recherche hébergés dans la base de données développée par la mission GAMeR, permettraient de construire rapidement des trames de données utilisables dans le cadre d'analyses statistiques des données.

II.1.b. *Travaux et financement envisagés*

Parmi les deux propositions précédemment citées (chap. III § II.1.a), il serait judicieux d'opter pour l'intégration de métadonnées d'échantillons dans la base de données développée par la mission GAMeR sous le format MongoDB. Effectivement, cette dernière est déjà unifiée et fonctionnelle sous ce format (Figure 19). Outre le développement de champs supplémentaires dans cette base de données gérée par la mission GAMeR, il sera aussi nécessaire de faire évoluer les procédures de soumission vers un transfert des métadonnées des échantillons en plus des contrôles d'ADN déjà en place (Annexe 13). Il est envisagé de faire financer ce projet AcSaMe par les plateformes SCA, ESV et/ou ESA (chap. III § II.1.a) nouvellement créés et supportés par les autorités françaises en vue d'implémenter la génomique pour la surveillance (Annexe 4).

Effectivement, la centralisation des métadonnées d'échantillons et des données génomiques rentrent dans le cadre des activités de ces plateformes d'épidémiologie (SCA, ESV et ESA) ayant pour objectif « *d'optimiser les actions et les coûts de la surveillance, par un partage de ressources, de compétences et d'outils* » (Annexe 4).

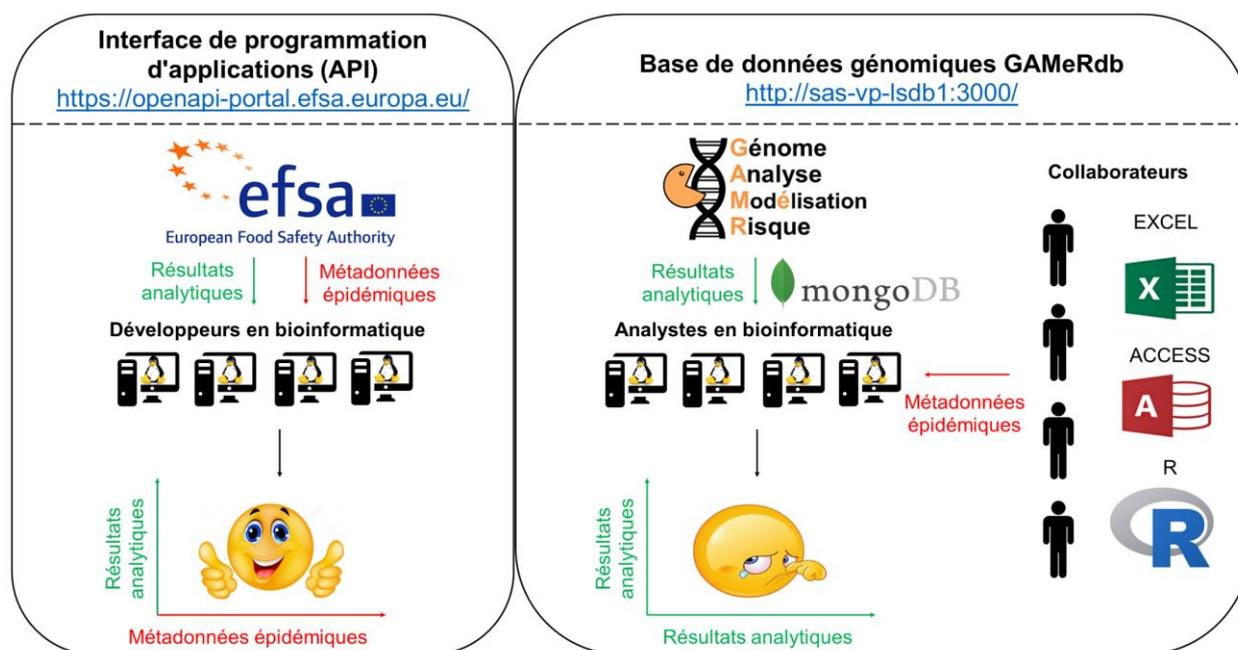


Figure 19 : Interface de programmation (API) de l'EFSA fournissant des métadonnées et résultats analytiques aux développeurs en bioinformatique versus base de données génomiques gérée par la mission génome analyse modélisation risque (GAMeR) qui n'intègre pas actuellement de métadonnées épidémiologiques pour les analystes en bioinformatique

II.2. Projet de référence : Évaluation des déterminants de non-diminution de l'incidence de *Salmonella* Enteritidis dans le cadre de l'application de la génomique aux filières aviaires (Projet ADONIS)

Sous le nom de « assessing determinants of the non-decreasing incidence of *Salmonella* Enteritidis » (ADONIS), le projet de référence ADONIS vise à comprendre quels sont les déterminants génomiques expliquant l'augmentation de la prévalence du sérovar *Salmonella* Enteritidis dans les filières des poules pondeuses en Europe.

II.2.a. Contexte scientifique

D'après le rapport sur les zoonoses en 2016, la tendance à la baisse des cas de salmonellose dans l'union européenne (UE) a atteint un seuil [89]. L'EFSA et l'ECDC s'accordent à dire que depuis 2014 les cas infectieux de *Salmonella* Enteritidis chez l'homme dans l'UE ont augmenté de 3% et que la prévalence chez les poules pondeuses est passée à cette période de 0,7% à 1,21% [89]. En 2016, *Salmonella* Enteritidis représentait 59% des cas de salmonelloses contractées dans l'UE sur un total de 94 530 cas humains de salmonelloses [89], probablement associées à la consommation d'œufs, de produits à base d'œufs et de viandes de volailles [137, 138]. L'augmentation depuis 2014 de la prévalence de *Salmonella* Enteritidis dans les filières des poules pondeuses dans l'UE (Figure 20) [139], pourrait être en lien avec des variations génétiques des souches responsables de l'expansion clonale d'une lignée présentant par exemple des aptitudes à la colonisation, la réplication, la résistance et/ou la virulence.

II.2.b. Travaux et financement envisagés

Financé par les états membres de l'EU et piloté par le « rijksinstituut voor volksgezondheid en milieu » (RIVM), le projet ADONIS est organisé en plusieurs tâches, dont l'une en génomique est à la charge du « statens serum institut » (SSI) en collaboration avec l'ANSES et le « państwowy instytut weterynaryjny » (PIWET). Dans le cadre de la tâche en génomique, j'ai soumis avec mon responsable Michel-Yves Mistou un projet en 2019 visant à appliquer l'outil microbial-GWAS récemment développé par la mission GAMeR [102] sur la base de nos compétences dans ce domaine [99, 128]. L'objectif associé serait d'identifier s'il existe des gènes accessoires ou variants du coregénom (SNPs et InDels) pouvant expliquer l'augmentation depuis 2014 de la prévalence de *Salmonella* Enteritidis dans les filières de poules pondeuses dans l'UE (Figure 20) [139]. Les ovo produits étant fortement suspectés d'être à l'origine [137, 138] des cas humains de salmonelloses à *Salmonella* Enteritidis dans l'UE [89], je propose de conduire ces travaux sur une collection d'environ 400 génomes provenant des collections du LSAL et ou des archives internationales en libre accès de données NGS. Le recrutement d'un(e) étudiant(e) en master financé(e) par le LSAL pour développer l'environnement Conda de l'outil microbial-GWAS [102], a été programmé, ainsi que le financement d'un(e) collègue postdoctorat pour supporter le projet ADONIS.

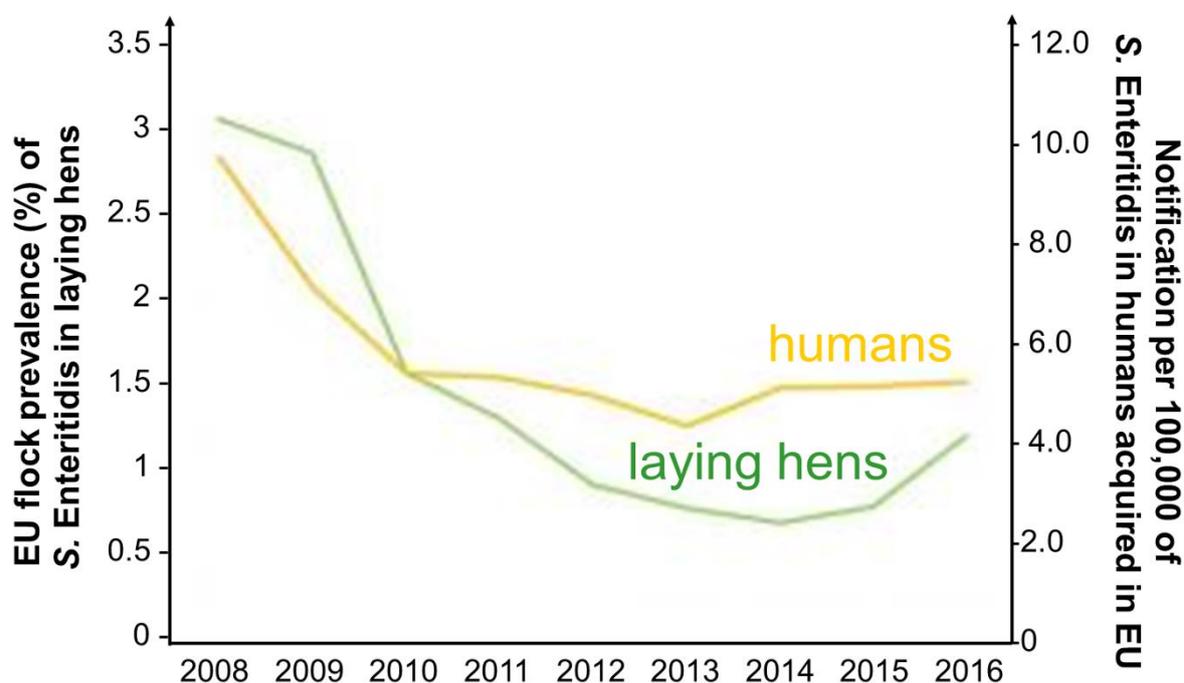


Figure 20 : Évolution dans l'union européenne (2008 et 2016) de la prévalence dans la filière des poules pondeuses (%) et des notification de cas humains (pour 100 000) de *Salmonella enterica* spp. *enterica* sérovar Enteritidis [89]

II.3. Projet de recherche : Évaluation à l'échelle génomique des sources hydriques à l'origine des TIAC à *Salmonella* dans l'union européenne (Projet FoWaBoSa)

Le projet de recherche « food and water-borne *Salmonella* » (FoWaBoSa) compte comparer les proximités génétiques des génomes de *Salmonella* issus de l'eau de surface, en particulier l'eau d'irrigation, et d'aliments potentiellement contaminés par cette dernière en Europe ou en France.

II.3.a. Contexte scientifique

II.3.a.i. Pathogène alimentaire en Europe

Salmonella est une bactérie intestinale et comporte des sérovars présentant des profils d'adaptation aux animaux multi- (généralistes) ou mono-hôtes (spécialistes). Il a

déjà été démontré que le serovar multi-hôtes Enteritidis a donné naissance aux lignées aviaires Gallinarum et Pullorum [101], et que le serovar multi-hôtes Typhimurium a conduit à l'apparition de lignées contemporaines spécialisées pouvant s'implanter dans des hôtes d'élevages [140]. Les *Salmonella* généralistes semblent donc être à l'origine de l'apparition de lignées spécialisées responsables de pathologies animales. Par ailleurs, *Salmonella* est actuellement la première cause de cas infectieux groupés humains d'origine alimentaire en France, ainsi que la seconde cause des épidémies humaines d'origine zoonotique dans l'UE en 2015 [88], 2016 [89] et 2017 [87], dernière *Campylobacter*. En Europe, les épidémies humaines d'origine hydrique (c.à.d. eau potable du robinet ou de puits) ont majoritairement été attribuées à des norovirus, d'autres calciviruses, des *E. coli* productrices de Shiga toxines ou des *Campylobacter* en 2015 [88], 2016 [89] et 2017 [87]. Néanmoins, aucune épidémie humaine causée par *Salmonella* n'a été associée officiellement à des origines hydriques entre 2015 et 2017 (Figure 21).

II.3.a.ii. Pathogène alimentaire d'origine hydrique aux États-Unis

D'après le « centers for disease control and prevention » (CDC), 53% des TIAC aux États-Unis ont été causées par *Salmonella* de 2006 à 2017 [141]. Environ 32% de ces dernières ont été associées à la consommation de denrées contaminées [141]. Les investigations associées suggèrent que l'eau d'irrigation est à la fois une source de contamination des denrées et un véhicule de transport de cette bactérie pathogène [141]. Après l'excrétion intestinale par un hôte animal, *Salmonella* peut persister dans l'eau de rivière, de lacs et d'étangs, qui sont des sources majeures d'irrigation [141]. *Salmonella* peut aussi s'internaliser dans un hôte secondaire végétal comme la tomate, contourner son système immunitaire et le coloniser [142]. Aux États-Unis, plusieurs TIAC ont effectivement été attribuées aux sérovars Newport dans des tomates (2005) [143] et SaintPaul dans des poivrons (2008) [144] contaminés par l'eau d'étangs utilisés pour l'irrigation. D'autres TIAC à *Salmonella* aux États-Unis ont été soupçonnées d'avoir été causées par des souches des sérovars Thompson (2010, 2012, 2013), Enteritidis (2013) ou Javiana (2012) présentant des profils génétiques identiques à des isolats d'eau de surface d'étangs [145] (Figure 21).

II.3.a.iii. Origine hydrique potentielle des TIAC en Europe ou en France

L'absence de TIAC humaines à *Salmonella* attribuées à des origines hydriques en Europe entre 2015 et 2017 [87–89], ainsi que la multitude de TIAC aux États-Unis attribuées à des sérovars issus de l'eau d'irrigation [143–145], peuvent refléter des différences dans les chaînes de production agroalimentaires et/ou dans les systèmes d'investigations des TIAC (Figure 21). Bien que les études en génomique sur les *Salmonella* d'origine animale [101, 102] et alimentaire [96, 98, 100] soient nombreuses, les rares études sur les *Salmonella* d'origine hydrique n'intègrent pas de génomique [146]. Sur la base des compétences de la mission GAMeR en reconstruction phylogénomique [96, 97, 99, 101, 102, 104, 118, 120, 128] et assignation de génomes à des événements épidémiques [92], nous souhaitons mettre en lumière quelles sont les proximités génétiques des génomes de *Salmonella* issus de l'eau potable (c.à.d. eau du robinet, eau de puits), d'eau de surface (c.à.d. rivières, lacs, étangs) et d'aliments potentiellement contaminés par cette dernière en Europe ou en France.

II.3.b. Travaux et financement envisagés

Sous le pilotage de la mission GAMeR, le projet FoWaBoSa est construit en collaboration avec les laboratoires Eau de Paris, le Laboratoire d'hydrologie de Nancy (LHN) qui est le LNR « eaux destinées à la consommation humaine, eaux minérales naturelles et eaux de loisirs - biologie », ainsi que l'unité SEL du LSAL qui gère une base de données dédiée à *Salmonella* à travers l'outil : « application pour la centralisation et le transfert de données dédiées à l'épidémiosurveillance opérationnelle des laboratoires » (ACTEOLab) [147]. À date, cette base ACTEOLab comporte des milliers

de souches d'origines alimentaire et animale, 60 souches de *Salmonella* isolées d'alimentation hydrique humaine et environ 5 000 souches d'écosystèmes hydriques. En discussion actuellement, la sélection des isolats d'intérêt se fera sur la base des collaborations des laboratoires Eau de Paris, du LHN en tant que LNR et du LSAL, afin de construire un plan d'expérience équilibré sur les variables zones géographique, origine du prélèvement et date d'isolement. Une attention particulière sera aussi apportée aux archives internationales en libre accès de données NGS, comme par exemple Enterobase qui comporte 226 231 enregistrements de *Salmonella* en date du 20 août 2019, dont 22 562 présentent des données NGS attachées. La soumission de ce projet FoWaBoSa est envisagée aux appels à financement ANSES de sujet de doctorat avec un laboratoire extérieur à l'ANSES.

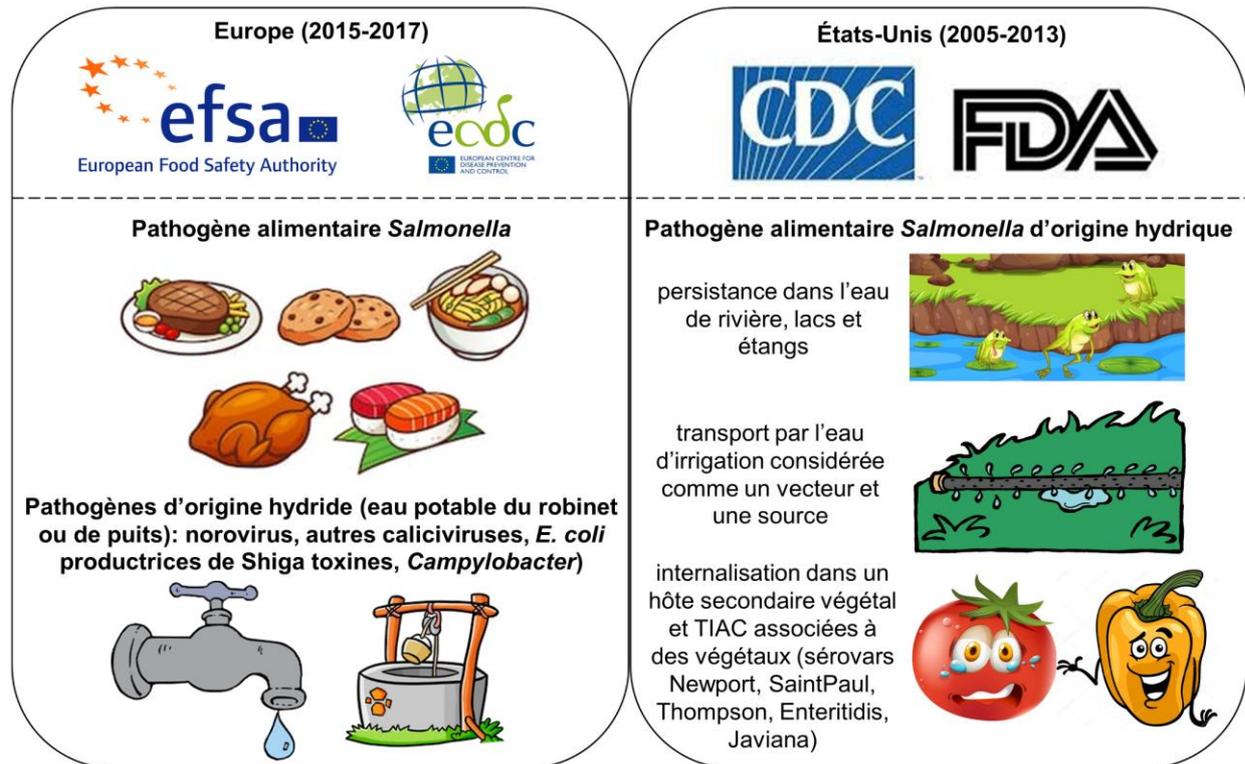


Figure 21 : Situations contrastées entre le pathogène alimentaire *Salmonella* en Europe (2015-2017) et le pathogène alimentaire *Salmonella* d'origine hydrique aux États-Unis (2005-2013)

*« Un expert est une personne qui a fait toutes les erreurs
qui peuvent être faites dans un domaine très étroit »*

Niels Bohr

DISCUSSION

En lien avec les objectifs qui lui ont été confiés, la mission GAMeR (chap. I § III), a initié des développements analytiques en génomique (Tableau 3) et la construction d'un réseau Linux de traitement de données génomiques (Figure 10) depuis 2014 en appui pour le LSAL (chap. II §II). La présente discussion concerne les investissements en bioinformatique (disc. § I), ainsi que l'accessibilité aux métadonnées épidémiques (disc. § II), nécessaires à l'harmonisation de la génomique des pathogènes d'origine alimentaire (disc. § III) dans un soucis de faible impact environnemental (disc. § IV).

I. Investissements en capacités de stockage, ressources de calcul et compétences en bioinformatique

Sans prendre en compte les caractéristiques des processeurs, les capacités de stockage et les ressources de calcul de l'institut anglais PHE (intro. § II.4 : stockage de 1 700 TB et calcul sur 656 cores distribués en 34 nœuds avec 5,4 TB de RAM), sont respectivement 23 et 2-4 fois supérieures à celles de la mission GAMeR (chap. II § II.1.b : stockage de 74 TB et calcul sur 340 cores distribués en 7 nœuds avec 2,4 TB de RAM). Ces différentiels de capacités de stockage et ressources de calcul expliquent donc en partie que sur une période similaire de 6 ans, incluant environ 2 ans de développements et 4 ans d'application, PHE a pu atteindre un total de 52 590 génomes assemblés (intro. § II.4). Par apposition, la mission GAMeR a réalisé 6 339 génomes assemblés en 5 ans d'existence (Figure 11). L'augmentation du nombre de génomes traités par la mission GAMeR pourra donc se faire à la condition *sine qua non* d'investir dans les capacités de stockage et ressources de calcul, avec une priorité pour le stockage qui est actuellement saturé à hauteur de 89% en date du 31 juillet 2019. Tout comme PHE, l'augmentation de la charge salariale dédiée aux développeurs informatiques dans la mission GAMeR (chap. I § III.1) est une étape nécessaire à l'implémentation de l'analyse de données en génomique bactérienne (intro. § II.4). Dans la sphère des entreprises privées, Chan Zuckerberg, la femme du fondateur de Facebook, développe actuellement une société BioHub visant à combattre les maladies infectieuses à l'échelle mondiale (Adresse web 8). Cette société BioHub ne dévoile pas son plan d'investissement en matériel, mais pourra assurément compter sur le savoir-faire en gestion de données massives de la société Facebook. Les sphères publiques et privées ont donc le même objectif d'implémenter la génomique et devraient unir leurs ressources bioinformatiques.

II. Accessibilité aux métadonnées épidémiques pour les développeurs et analystes en bioinformatique

L'EFSA a dernièrement rendu accessibles ses métadonnées aux développeurs par la mise en place d'une API (Adresse web 7). Pour remédier à l'absence d'accès aux métadonnées épidémiques des échantillons traités par la mission GAMeR (chap. III § II.1), il serait d'un intérêt majeur d'intégrer ces dernières à la base de données génomiques développée par la mission GAMeR (Figure 10). Cette accessibilité aux métadonnées épidémiques du LSAL permettrait d'envisager des développements d'interfaces graphiques reliées à ces dernières (Figure 19). D'ailleurs, l'EFSA et l'ECDC proposent aujourd'hui des outils de visualisation des données relatives aux pesticides (Adresse web 9) et aux résistances aux antibiotiques (Adresse web 10). L'adjonction des métadonnées épidémiques à la base de données GAMeR (Figure 10), notamment géographique, permettrait aussi d'envisager le développement d'outils de visualisation des prévalences des génotypes des genres bactériens à prendre en considération en sécurité alimentaire. Ceci a d'ailleurs été proposé par le CDC (Adresse web 11) et Santé Publique France (Adresse web 12) qui ont mis à disposition leurs outils de visualisation des données épidémiologiques.

III. Harmonisation de la génomique des pathogènes alimentaires pour l'efficacité de la santé publique

La diversité des développements analytiques en génomique des pathogènes d'origine alimentaire, à la fois dans la mission GAMeR (Tableau 3), mais aussi à un niveau international (Tableau 1), est un aspect très positif concernant l'augmentation des connaissances. Néanmoins, les chevauchements de certains développements de différentes équipes comme notamment les étapes d'assemblage et d'identification de variants (Tableau 1 et Tableau 3) devraient être évités afin de minimiser le temps de développement d'outils analytiques existants. À côté des comparaisons des outils analytiques coordonnées par l'EFSA et l'ECDC [80], les initiatives en génomique des états membres devraient s'orienter vers le développement d'outils présentant des objectifs différents, afin de les combiner et améliorer plus efficacement le contrôle des pathogènes alimentaires.

IV. Impact environnemental majeur de la génomique

Tout comme nous l'observons à l'échelle de la mission GAMeR (Tableau 4), même si les coûts du séquençage ont diminué drastiquement (intro. § II.1), les tailles des bases de données internationales d'archives NGS augmentent inexorablement (intro. § II.2). Ceci engendre un déplacement de coûts vers les capacités de stockage, les ressources de calcul (disc. § I) et la charge salariale de développeurs et analystes en bioinformatique (disc. § II). À l'heure de l'accord de Paris relatif au climat et au réchauffement climatique, l'augmentation exponentielle des capacités de stockage des données génomiques pose donc question quant à sa pérennité écologique. Effectivement, les centres de données (« data center ») accueillant ces dernières ont une empreinte écologique non négligeable, même si des études de réutilisation des calories produites sont en cours [148].

CONCLUSION

Comme demandé par le laboratoire qui m'accueille actuellement et stipulé dans les textes de loi afférents, le présent mémoire décrit respectivement la mission GAMeR du LSAL de l'ANSES et mon dossier de travaux accompagnés d'une synthèse de mes activités scientifiques, afin de mettre en exergue ma capacité à animer la recherche. Outre l'exercice de compilation de mes travaux passés et présents, la rédaction du présent mémoire visant à candidater à l'obtention de l'HDR, m'a aussi permis d'accentuer mon regard critique sur mes propres activités de recherche. Ceci m'a permis de construire un programme de recherche visant à poursuivre ma participation à l'intégration de la génomique des bactéries pathogènes alimentaires pour la santé publique moderne en France, Europe et dans le monde. Mes activités de recherche sont internationales entre la France, les États-Unis et le Canada, ainsi qu'originales en raison de l'innovation des procédures analytiques les soutenant. Ces activités de recherche sont aussi larges au regard de leurs domaines d'application (c.à.d. médical, environnemental et alimentaire) et de leurs disciplines scientifiques (c.à.d. bactériologie, biologie moléculaire et génomique). Ces travaux sont également communautaires au vu des collaborations avec d'autres chercheurs et étudiants de tous niveaux. En conséquence, mes activités de recherche ont participé et continueront à contribuer à l'amélioration de la santé publique. Je souhaite dans le futur poursuivre mes travaux à l'ANSES, participer davantage à la vie de l'ED ABIES qui intègre mes champs disciplinaires, ainsi que continuer à encadrer des thèses de doctorat en tant que directeur de thèse et être impliqué dans des jurys de thèse en tant que rapporteur. Ce présent mémoire me permet de soumettre ma candidature à l'obtention de l'HDR sous l'égide de l'UPE, représentant pour moi l'instance qui m'a permis d'obtenir mon doctorat.

« *Knowledge of languages is the doorway to wisdom* »

Roger Bacon

SYNTHÈSE EN ANGLAIS

The present manuscript entitled “The bacterial genomics in support of public health” supports my candidacy to obtain the last French diploma called certification to manage research (“habilitation à diriger des recherches: HDR”). Under the agreements of the University Paris-Est and jury members, this HDR candidacy must be accompanied by a HDR defence. More precisely, my HDR candidacy is based on the present manuscript written in French describing my detailed CV (Annexe 1 et Adresse web 2), current position (CHAPITRE I : LIEU D’ACCUEIL), research career (CHAPITRE II : BILAN DE CARRIÈRE) and future projects in bacterial genomics (CHAPITRE III : PROGRAMME DE RECHERCHE), as well as my published articles written in English (Adresse web 3). My detailed CV is also accessible online in English (Adresse web 13) and I plan to perform my HDR defence in English according to the supervisors of my HDR candidacy, Anne Brisabois (Anses) and Michel-Yves Mistou (INRA). This synthesis of the present manuscript will describe my host laboratory (synth. § I), career outcome (synth. § II) and future projects (synth. § III).

I. Host laboratory

Within the framework of the “French Agency for Food, Environmental and Occupational Health & Safety” (ANSES) (synth. § I.1), I am in charge of the implementation of “next generation sequencing” (NGS) and “whole genome sequencing” (WGS) tools for the laboratory for food safety (LSAL) (synth. § I.2) in the context of the team “Genome Analysis Modelling Risk” (GAMeR) (synth. § I.3).

I.1. French Agency for Food, Environmental and Occupational Health and Safety

ANSES was created on 1 July 2010 and is an administrative public establishment accountable to the French Ministries of Health, Agriculture, the Environment, Labour and Consumer Affairs. The government, stakeholders represented on its Board of Administrators and labour unions can request to ANSES in the fields of environmental health, occupational health, food and nutrition, animal health and well-being and plant health (Annexe 3). The ANSES Board of Administrators integrates the five colleges of the Grenelle Environment round table, namely public authorities, social partners, professional bodies, non-governmental organization and non-profit associations, as well as elected officials. ANSES is organized in 9 reference and research laboratories located in France close to the various sectors associated with their activities, including animal health and well-being, chemical and biological food safety, as well as plant health (Figure 4).

I.2. Laboratory for food safety

Specialised in the biological (bacteria, viruses, parasites) and chemical hazards (toxins, metals, pesticides, organic pollutants, histamine) which have an effect on food safety and quality, the LSAL is located in Maisons-Alfort and Boulogne-sur-Mer. The Maisons-Alfort site hosts 120 professionals in food safety and the Boulogne-sur-Mer site includes members involved in the quality and hygiene of fishery and aquaculture products. The LSAL fulfills ANSES’s missions based on activities related to surveillance, reference and research (Figure 5). The LSAL is organized according to mandates of European Union Reference Laboratory and National Reference Laboratory (Annexe 5). More precisely, the LSAL develops methods for detecting, characterising and measuring of hazards found in foods, analyses their causes and development factors in order to monitor them and report the emergence or re-emergence of organisms and compounds.

I.3. Genome Analysis Modelling Risk

The GAMeR team was created non-officially in 2014 and officially in 2016 by Michel-Yves Mistou in order to support the LSAL activities in bacterial genomics (Annexe 6).

Since 2015, I am in charge of the implementation of WGS tools, data and database management within the transversal GAMeR team of the LSAL at ANSES (Figure 6). My activities focus mainly on the implementation of (i) NGS exploited by the different units of LSAL, (ii) open-source Linux networking systems, (iii) NoSQL-oriented database, and (iv) post-NGS data processing through in-house developed workflows for variant calling, de novo assembly, WGS-based subtyping (i.e. resistome, plasmidome) and statistical analysis for outbreak investigation.

II. Career outcome

My career outcome includes 27 studies constituted of my PhD manuscript obtained the 28th February 2011 [115], submissions of patents in 2018 [91], a book chapter [116], 2 international publications where my contributions were acknowledged [96, 117], as well as 20 international publications published as co-auteur or main authors [92, 94, 95, 97–102, 118–128] and 2 under peer-reviewing [103, 104]. Excepted my PhD manuscript [115], the submissions of patents [91], the book chapter [116] and the acknowledgements [96, 117], these publications represent my contributions as last author [101, 102], first author [92, 118, 120, 122–127] and intermediate author [94, 95, 97–100, 103, 104, 119, 121, 128]. All these international publications are accessible online (Adresse web 3) and refer to my research studies during my Master degree [117, 127], PhD project [116, 122, 124–126], postdoctoral fellows [118–121, 123] and current position [92, 94–104, 128]. My open researcher and contributor identifier (ORCID) is 0000-0002-7480-4197 (Adresse web 4). Since 2007, I am involved in NGS, WGS and microbial genomics in collaborations with French, American, Canadian and European stakeholders (Tableau 2) focusing on risk assessment related to human (synth. § II.1) and animal (synth. § II.2) health, as well as environmental (synth. § II.3) and food (synth. § II.4) safety.

II.1. Human health

I have been involved in human health research project during postdoctoral fellows between 2011 and 2015, funded by the “Fonds de recherche du Québec - Santé” (FRSQ), in the international scientific environment of the Research Institute of McGill University Health Center (RI MUHC). I wrote three applications with my supervisor to obtain the corresponding funding (grants #29836 and #26274). My main objectives were to develop bioinformatics and molecular tools for Crohn’s disease diagnosis [123], investigation of tuberculosis (TB) outbreaks [120] and population genomics studies [118] of TB in Inuit communities (Canada). I collected human samples, managed and validated large-scale genomic data (n=200, Miseq and PacBio) and statistically analysed the evolution and impact of virulence factors by *in vivo*, *in vitro* [121] and *in silico* [119] models. The main results from these postdoctoral fellows in Canada [118–121, 123] showed that the *Mycobacterium tuberculosis* causing outbreaks in Inuit communities between 1990 and 2000 were introduced in the early 20th century from Europe, perhaps European whalers and explorers [118]. Because of a genomic profile lacking features of hypervirulence and antibiotic resistance, we concluded that the ongoing TB outbreaks were likely due to host or social factors conducive to transmission [118]. This collaborative postdoctoral studies in Canada have been presented to different stakeholders from the scientific communities, Inuit leaders, as well as provincial and territorial partners (Adresse web 13).

II.2. Animal health

I performed research studies related to animal health in a context of bovine tuberculosis suspicion in livestock and lack of understanding of transmission mechanisms of the environmental pathogenic bacteria *Mycobacterium avium* between human and animal. These research studies related to animal health were carried out during my Master degree in the bacterial zoonoses unit (UZB) of the laboratory for animal health (LSAN) at ANSES and focused on bacterial phylogenetic reconstructions and adaptation to animal hosts of *M. avium* by restriction fragment length polymorphism (RFLP) and variable number tandem repeat (VNTR) [127]. In addition to my analytic and writing

contributions, I participated to the conception and funding application to the « veterinary network of laboratories researching into improved diagnosis and epidemiology of mycobacterial diseases » (VENoMYC) in collaboration with the French agricultural research institute (INRA) (grant AFSSA-INRA #AIP P00297). The results of these research studies showed that the subspecies of *M. avium* were adapted to specific hosts including human, swine, bovine, avian, as well as wild and domestic animals [127].

II.3. Environmental safety

I performed my PhD project between 2007 and 2011 in collaboration with laboratories from Paris-Est University (ParisTech) and the Virginia polytechnic institute (VirginiaTech) in a context of surface and deep water restauration before 2015 (2000/60/CE) and effluent control (91/271/CE et 98/15/CE). During this project related to environmental safety, I conceptualized and wrote purpose of my PhD project with my supervisors to obtain the corresponding funding (grants PIREN-Seine, OPUR and CRECEP-STEAM). This PhD project aimed at collecting evidence of non-tuberculous mycobacteria (NTM) in watersheds by bacteriology [126], molecular biology [125] and genomics [122]. I collected samples from surface water, drinking water and soil, as well as recovered data on microbial hazards (*Mycobacterium*, *Escherichia* and *Enterococcus*) and concentrations of chemical parameters (metal, organic matter, pH, ions and conductivity) in order to statistically compare their behaviours in watersheds [124]. I also collected and compared molecular sequences and assembled genomes (n=28 from NCBI) to provide specific and sensitive markers of *Mycobacterium* to control authorities [122]. During this PhD in France and USA [116], I provided innovative microbiological risk assessment analytical tools and results to the laboratory controlling water distribution systems of Paris (Eau de Paris), wastewater treatment plants of the greater Paris Sanitation Authority (SIAAP) and the epidemiology team of the national center for mycobacteria (Public assistance Hospital of Paris: APHP). Based on molecular and genomics tools, microbiological risk assessment of watersheds in Paris allowed to identify main sources of mycobacteria [115].

II.4. Food safety

Within the objectives of my current position in the GAMeR team in LSAL (synth. § I.3), we collect, analyse and interpret large amounts of genomic data (synth. § II.4.a) through developing and documenting software (synth. § II.4.b) for consistent metadata and genomic data processing, storage and management (synth. § II.4.c) to ensure data format and standard outputs harmonization for research projects in collaboration with LSAL units (synth. § II.4.d) and projects coordinated by the GAMeR team (synth. § II.4.e).

II.4.a. Funding and human resources

With the support of my supervisor Michel-Yves Mistou between 2015 and 2019, I coordinated the setting up of the GAMeR team, screening and selecting candidates (i.e. from 2 to 4 permanent scientists in 3 years). I applied for research funding (COMPARE grant #643476, ANSES-INRA grant #Typautobac, EJP ADONIS grant #ADONIS and EJP RaDAR grant#773830) and I manage the development by chairing internal meetings, monitoring projects and tasks progress and employing annually between 2 and 4 students in master, PhD and postdoctoral fellowships (Figure 17).

II.4.b. Main genomic workflows

In collaboration with my GAMeR team (permanent colleagues Arnaud Felten and Ludovic Mallet) and working groups from LSAL in ANSES, I prepared procurements for acquisition of NGS data since 2015 (Tableau 4), as well as guidance documents to support LSAL colleagues and projects' partners on the submission of genomics data to the GAMeR curated database (Annexe 13). In the collaborative and international scientific ecosystem of the LSAL in ANSES, the GAMeR team ensures harmonization of data and

standard outputs of bacterial genomics developing and applying standardized workflows for variant calling (i.e. the tool iVARcall2 : independent variant calling analysis) (Figure 8) and *de novo* assembly (i.e. the tool ARTwork : assembly of reads and typing workflow) (Figure 9). I collaborate also with other LSAL units to support WGS analyses through guidance documents related to the 'wet' and 'dry' lab. I also participated to set up the procedures for automated DNA extraction for NGS and implementation of a MiSeq sequencer in LSAL, in addition to managing three proficiency tests organized by the Global Microbial Identifier (GMI) between 2015 and 2017.

II.4.c. Computing and storage system

The GAMeR team contributed to the development of a structured process of retrieving and combining data (i.e. metadata and genomic data) from different sources (e.g. local and public repositories) into valuable and meaningful information, ensuring quality, integrity and security of data flows and content. The GAMeR team also manage domain users and synchronisation of our Linux system with the Windows system of ANSES. With the support of the IT service of ANSES (Thomas Texier et Pierre-Yves Letournel), NGS data produced by the Brain and Spine Institute (ICM) (Yannick Marie) and scientific exchanges with the national genomic platform in ANSES (Yannick Blanchard), the GAMeR team built a backup system of our storage (74 TB used at 78-87% in February 2020) and computing systems (340 cores distributed through seven nodes with 2.4 TB RAM) (Figure 10) (Ludovic Mallet). Since 2015, the GAMeR team has managed the WGS (mainly illumina MiSeq, NextSeq and HiSeq) of around 1,000 bacterial strains per year for a total of 6,000 bacterial genomes collected today in contexts of European projects (Figure 11). Today, the GAMeR team is able to perform variant calling (i.e. the tool iVARcall2) (Figure 8) and *de novo* assembly (i.e. the tool ARTwork) (Figure 9) for 96 bacterial genomes in 400 minutes. All this pre-processed genomic data is available to the LSAL colleagues through the Linux network and MongoDB database (Figure 11) or Windows graphical interface developed by the GAMeR team (internal ANSES address: Adresse web 5) (Arnaud Felten, Kévin Durimel and Pauline Barbet). With this pre-processed genomic data (internal ANSES address: Adresse web 6 on Windows or « ssh -X ansesID@sas-vp-lsgw1 » on Linux) the LSAL users can perform more advanced analyses with the Linux network, the national Galaxy platform or other international genomics platforms (Tableau 1). Through methodological and guidance documents, the GAMeR team provides also training annually related to this Linux network to stakeholders involved in data analysis and interpretation.

II.4.d. Collaborative research projects

In partnership with PhD students and postdoctoral fellows, I have been involved in microbiological risk assessment and manuscript writing for publications concerning collaborative projects related to *Salmonella* serovar Derby in French avian and pig sectors (Yann Sévellec) [96, 98, 100], *Clostridium perfringens* outbreaks in France *Clostridium* (Abdelrahim Abakabir Mahamat) [97], *Listeria monocytogenes* persistence in food (Lena Fritsch et Clémentine Henri) [94, 99] and seafood processing plants in France (Federica Palma) [128], as well as *Bacillus thuringiensis* used as biopesticides in the French market (Mathilde Bonis) [104]. In addition to these collaborative projects with PhD students and postdoctoral fellows, I also contributed to other collaborative projects with colleagues from other LSAL units (Sabrina Cadel-Six and Patrick Fach). These collaborative projects allowed development of workflows aiming to estimate accuracy of molecular targets at the genomic level (i.e. the tool GTEvaluator) [95] (Annexe 7) and support patent applications related to specific and sensitive kmer detection (i.e. the tool GTFinder) [91] (Annexe 8). The GAMeR is currently involved in the development of other tools related to bacterial genomics (Tableau 3).

II.4.e. GAMeR research projects

Keeping up-to-date about WGS, I had a prominent role in ANSES as project leader in developing and providing innovative approaches to assess microbiological risks related to (i) outbreak investigations (i.e. the tool Pairwise-FBO) [92] (Figure 12 and Annexe 9), (ii) identification of coregenome fixed variants and metabolic pathways (i.e. the tool PhyloFixedVar) [101] (Figure 13 and Annexe 10), as well as (iii) association studies at the pangenomic scale (i.e. the tool microbial-GWAS) [102] (Figure 14 and Annexe 11). Firstly (i), I coordinated the COMPARE-funded project on outbreak investigation benchmarking three non-parametric statistical tests (i.e. Kolmogorov-Smirnov, Wilcoxon rank sum and Kruskal-Wallis) on six genomic features (i.e. SNPs with or without recombination, genes, kmers, cgMLST and wgMLST alleles) to discriminate outbreak cases from non-outbreak cases in collaboration with the *Salmonella* LSAL team and the National Reference Center of *Salmonella* (i.e. Pasteur institute) [92]. Secondly (ii), I coordinated a study which highlighted that adaptation as *Salmonella* serovars to animal hosts was mainly driven by coregenome fixed variants related to the amino acid metabolism [101]. Thirdly (iii), I supervised a PhD student (Méryl Vila Nova) in collaboration with INRA based on Genome Wide Association Study (GWAS) and Gene Ontology Enrichment Analysis (GOEA) workflows that the GAMeR team developed [102]. We showed in this study that the genetic and metabolic determinants of *Salmonella* adaptation to animal sources have been driven by the natural feeding environment of the animal, distinct livestock diets and environmental stimuli [102]. Animal-specific physiological properties and anthropological activities for the livestock health protection drove also *Salmonella* adaptation to animal sources (Figure 15 and Annexe 12) [102]. Outputs from these studies, can be used to feed source attribution models in order to assess *Salmonella* risk through the food chain [102]. With publication objectives, I continue currently to work for food safety focusing on antimicrobial resistance (AMR) related plasmids (EJP RaDAR grant #773830) and genetic features explaining *Salmonella* Enteritidis increasing in Europe since 2015 (EJP ADONIS grant #ADONIS).

III. Future projects

According to the LSAL organisation (synth. § I.2), I propose to perform in a near future projects in the fields of surveillance (synth. § III.1), reference (synth. § III.2) and research (synth. § III.3).

III.1. Surveillance project: Accessibility to LSAL epidemiological metadata and integration in the genomic database developed by the GAMeR team (Project AcSaMe)

Due to the related scientific context (synth. § III.1.a), I propose to the surveillance project « accessibility of sampling metadata » (AcSaMe) (synth. § III.1.b).

III.1.a. Scientific context

Since 2019, the European Food Safety Authority (EFSA) proposes an Application Programming Interface (API) to provide access to food safety data to developers (Adresse web 7). The implementation of a genomic database in LSAL by the GAMeR team, nicknamed GAMeRdb (Tableau 3 et « stockage bio » de la Figure 10), is currently not linked to epidemiological metadata concerning the samples submitted to the GAMeR team (Annexe 13). This situation slow down the application of tools developed by the GAMeR team (Tableau 3). Because the epidemiological metadata of the LSAL are not harmonized into a single format (i.e. R dataframes, Excel spreadsheets, Access data base, paper records), it seems suitable to records epidemiological metadata into the GAMeRdb in parallel to the NGS samples submitted for pre-processing (Figure 19). Indeed, the GAMeRdb is already harmonized into a single MongoDB format and linked to a computing cluster (Figure 10). This unified Linux network would allow requests of

genomic and epidemiological data at the same time in order to build instantaneously dataframe form advanced statistical analyses (e.g. Bash or Python languages).

III.1.b. Plan and funding

The project AcSaMe would include to create new fields in the GAMeRdb dedicated to epidemiological metadata (internal ANSES address: Adresse web 5) and to modify the procedures of NGS sample submission to the GAMeR team (Annexe 13). We plan to support this project AcSaMe with three epidemiological surveillance platforms in France, namely animal health epidemiological surveillance (ESA), sphere of action of the plant health epidemiological surveillance (ESV) and food chain surveillance platforms (SCA) (Annexe 4). The epidemiological surveillance platforms have to centralize epidemiological metadata and genomic data, as well as “optimize activities and the cost of surveillance by pooling resources, skills and tools” (Annexe 4).

III.2. Reference project: Assessing determinants of the non-decreasing incidence of *Salmonella* Enteritidis (Project ADONIS)

Due to the related scientific context (synth. § III.3.a), I propose to the reference project « assessing determinants of the non-decreasing incidence of *Salmonella* Enteritidis » (ADONIS) (synth. § III.3.b).

III.2.a. Scientific context

According to the report about zoonoses in 2016, the decreasing trend of salmonellosis cases is currently stable in European Union (EU) [89]. The EFSA and European Centre for Disease Prevention and Control (ECDC) emphasize that the infectious human cases of *Salmonella* Enteritidis in EU, increased of 3% since 2014 with prevalence in layer hens increasing from 0,7% à 1,21% at the same period [89]. In 2016, *Salmonella* Enteritidis represented 59% of salmonellosis cases in EU for a total of 94 530 human salmonellosis cases [89], probably associated to the consumption of eggs, egg-based products and poultry meats [137, 138]. The increasing of prevalence of *Salmonella* Enteritidis in layer hens from EU since 2014 (Figure 20) [139], may be explained by genetic mutations promoting the clonal expansion of lineages through aptitudes for colonization, replication, resistance and/or virulence.

III.2.b. Plan and funding

The project ADONIS is supported by the EU member states and managed by the “rijksinstituut voor volksgezondheid en milieu” (RIVM) and organized in different working packages (WP). Through a collaboration between ANSES, “statens serum institut” (SSI), “państwowy instytut weterynaryjny” (PIWET), I am more precisely involved in the working package 4 (WP4) dedicated to *Salmonella* genomics. Together with my supervisor Michel-Yves Mistou, we proposed a project in 2019 aiming to apply the tool Microbial-GWAS recently developed by the GAMeR team [102] based on our skills in this discipline [99, 128]. The objective is to identify mutations (i.e. genes, SNPs and InDels) associated with the increasing of incidence of *Salmonella* Enteritidis in layer hens in EU since 2014 (Figure 20) [139]. Because the egg-based products are suspected to be associated [137, 138] with the human salmonellosis cases in EU [89], I propose to lead the WP4 of the ADONIS project based on a collection of 400 genomes of *Salmonella* Enteritidis from LSAL collection and/or international archives. In the context of this WP4, we plan to recruit a Master student to package in conda the tool microbial-GWAS [102] and postdoctoral fellow to collect data and support the advanced genomic analyses.

III.3. Research project: Evaluation at the genomic scale of hydric sources causing *Salmonella* foodborne outbreaks in European Union (Project FoWaBoSa)

Due to the related scientific context (synth. § III.3.a), I propose to the reference project « food and water-borne *Salmonella* » (FoWaBoSa) (synth. § III.3.b).

III.3.a. Scientific context

The generalist serovars of *Salmonella* (i.e. multi-animal host) seems to be the origin of specialist serovars of *Salmonella* (i.e. mono-animal host) (e.g. Gallinarum and Pullorum from Enteritidis [101] of specific contemporary lineages from Typhimurium [140]). Otherwise, *Salmonella* is the first cause of grouped human infections from food origin in France, as well as the second cause of human outbreaks from zoonotic origins in EU in 2015 [88], 2016 [89] and 2017 [87], after *Campylobacter*. In Europe, the human outbreaks of hydric origins (i.e. tap water, water well) have mainly been associated with norovirus, other caliciviruses, Shiga-toxin-Producing *Escherichia coli* or des *Campylobacter* in 2015 [88], 2016 [89] and 2017 [87]. Nevertheless, human outbreaks caused by *Salmonella* have never been associated to hydric origins in EU between 2015 and 2017 (Figure 21). This situation is completely different in USA. According to the « centers for disease control and prevention » (CDC), 53% of foodborne outbreaks (FBOs) in USA have been caused by *Salmonella* from 2006 to 2017 [141]. Around 32% of these FBOs have been associated to consumption of contaminated foods [141]. The associated investigations suggest that irrigation water is a source and vehicle of pathogenic bacteria [141]. After animal excretion, *Salmonella* can persist in surface water, lakes and ponds, which are major sources of irrigation [141]. *Salmonella* can also internalize itself into secondary vegetable hosts like tomato and colonize it bypassing immune system [142]. In USA, several FBOs have been attributed to serovars Newport in tomato (2005) [143] and SaintPaul in pepper (2008) [144] contaminated by water from ponds used for irrigation. Because of similar genetic profiles, other *Salmonella* FBOs in USA have been suspected from hydric origins, especially for the serovars Thompson (2010, 2012, 2013), Enteritidis (2013) or Javiana (2012) [145] (Figure 21). Based on the aptitudes of the GAMeR team to reconstruct phylogenomic histories [96, 97, 99, 101, 102, 104, 118, 120, 128] and assign genomes to FBOs [92], we plan to estimate genetic proximities between *Salmonella* genomes from water for human consumption (i.e. tap water, well water), surface water (i.e. rivers, lakes, ponds) and food potentially contaminated by hydric origins in Europe or France.

III.3.b. Plan and funding

The GAMeR team plan to manage the project FoWaBoSa in collaboration with the laboratories Eau de Paris, the Nancy Laboratory for Hydrology as National Reference Laboratory for “water for human consumption, natural mineral water from springs and recreational water”, as well as the LSAL unit “*Salmonella* and *Listeria*) which is in charge of a *Salmonella* epidemiological database called ACTEOLab [147]. Today, this database ACTEOLab includes several thousands of strains from food and animals, 60 strains from water for human consumption and 5 000 strains from hydric ecosystems. We plan also to use the international archives of NGS data (e.g. 226 231 *Salmonella* records in Enterobase the 20th August 2019 including 22 562 records with associated NGS data) to build a genome dataset balanced in terms of geographical origins. We plan to submit this project FoWaBoSa to ANSES funding calls in collaboration with external laboratories.

« Si chaque découverte entraîne plus de questions qu'elle n'en résout, la part d'inconnu s'accroît à mesure que nos connaissances augmentent »

Pierre Vathomme

ANNEXES

Annexe 1 : Curriculum Vitae détaillé

Mes activités de recherche vont de la microbiologie conventionnelle aux techniques génomiques, et me permettent d'appliquer ces compétences en microbiologie alimentaire, évaluation des risques environnementaux, contrôle d'infections zoonotiques et santé publique.

ETAT CIVIL

Adresse personnelle : Appartement 44, 22 rue Eugène Renault, 94701 Maisons-Alfort (France).
Adresse professionnelle : ANSES-LSAL-GAMeR, 14 rue Pierre et Marie Curie 94701 Maisons-Alfort (France).
Contacts personnelles : +33 (0) 6 61 90 04 52, nicolas.radomski.phd@gmail.com.
Contacts professionnels : +33 (0) 1 49 77 27 53, nicolas.radomski@anses.fr.
Citoyenneté : Citoyen français, résident permanent au Canada, vivant en concubinage.
Langues : Français (maternelle), anglais (lu, écrit, parlé), allemand (lu).
Site internet : <http://www.nicolas-radomski.net>
ORCID : <https://orcid.org/0000-0002-7480-4197>



DISCIPLINES/SPÉCIALITÉ et COMPÉTANCE PARTICULIÈRES

Biologie cellulaire, biologie des populations et écologie, biologie des organismes.
 Microbiologie alimentaire, évaluation des risques environnementaux, contrôle d'infections zoonotiques.
 Microbiologie conventionnelle, biologie moléculaire, génomique bactérienne.
 Génomique bactérienne, statistique, bioinformatique.

FONCTION ACTUELLE

Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (ANSES).
 Laboratoire de sécurité des aliment (LSAL).
 Mission génome analyse modélisation risque (GAMeR).
 Chargé de projet recherche en génomique bactérienne.

PARCOURS PROFESSIONNEL

- 2015** **Chargé de projet recherche en génomique bactérienne** - Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (ANSES), Laboratoire de sécurité des aliment (LSAL) - Mission génome analyse modélisation risque (GAMeR) - Maisons-Alfort, France :
 Implémentation du séquençage haut débit, construction d'un réseau Linux de traitement et stockage des données, formation Linux aux équipes du laboratoire, génomique bactérienne (détection de variations génétiques, assemblage *de novo*, mécanismes moléculaires de virulence et d'adaptation aux hôtes, épidémiologie moléculaire, étude d'association, antibiorésistance, mobilome), recrutement et intégration d'étudiants et collègues tous les 6 mois - M.Y. Mistou (INRA-ANSES).
- 2015**
(1 mois) **Assistant en enseignement** - École des ponts ParisTech (ENPC) - Parcours d'ingénieur - Marne-la-Vallée, France :
 Théories en génomique bactérienne, utilisation de lignes de commandes sur Linux, paramétrage de programmes, exécution de programmes, organisation de tâches bioinformatiques successives - F.S. Lucas (LEESU).
- 2011-15**
(43 mois) **Postdoctorat** - Institut de recherche du centre de santé de l'université de McGill (RI MUHC) - Montréal QC, Canada :
 Installation de logiciels et programmes de gestion de laboratoire en analyses statistiques, analyses phylogéniques, comparaisons de séquences et de génomes, gestion de données de séquençage haut débit, détection moléculaire et quantification de pathogènes mycobactériens durant des infections expérimentales, développement et application de successions de programmes pour des projets de séquençage haut débit d'épidémiologie moléculaire de la tuberculose - M.A. Behr (RI MUHC).
- 2007-11**
(39 mois) **Doctorant** - École des ponts ParisTech (ENPC) - Laboratoire eau environnement et systèmes urbains UMR MA-102 (LEESU) - Marne-la-Vallée, France :
 Installation de laboratoires de bactériologie et biologie moléculaire, harmonisation des méthodes d'isolement mycobactérien dans des échantillons environnementaux, développement de méthodes de quantification par PCR en temps réel, génotypage d'isolats par multilocus sequencing analysis (MLST) et comparaison de génomes pour améliorer des cibles moléculaires - R. Moilleron, F.S. Lucas (LEESU).
- 2008-10** **Assistant en enseignement** - Licence analyse biologique et biochimique (ABB) et master microbiologie appliquée et génie biologique (MAGB) aux université Paris-Est de Créteil (IUT UPEC) et Paris Diderot (Paris-Sud) - Créteil et Orsay, France :

- (160 heures) Chargé de cours, travaux pratiques et travaux dirigés sur la classification, l'identification et la caractérisation de microorganismes pathogènes (*Enterobacteriaceae*, *Vibrionaceae*, *Pseudomonadaceae*, *Micrococciaceae*, *Streptococcaceae*, *Mycobacteriaceae*, *Neseriaceae*) pour des étudiants de licence et master - F. Odelin (IUT UPEC), C. Rousseau (IUT UPEC), C. Sola (Paris-Sud).
- 2010 (3 mois) **Stagiaire doctorant** - Institut polytechnique de Virginie - Département de biologie (VirginiaTech) - Blacksburg VA, USA) :
Étude des mycobactéries non-tuberculeuses dans un bassin versant de Virginie par PCR en temps réel, transfert d'une méthode de quantification par PCR en temps réelle des mycobactéries dans des échantillons environnementaux, rédaction d'un chapitre de livre - Falkinham III J.O. (VirginiaTech), R. Moilleron (LEESU), F.S. Lucas (LEESU).
- 2007 (6 mois) **Stagiaire en master** - Agence française de sécurité sanitaire des aliments (AFSSA) - Unité de zoonoses bactériennes (UZH) - Maisons-Alfort, France :
Diversité génétique de *Mycobacterium avium* d'origines humaine et animale par variable number of tandem repeats et mycobacterial interspersed repetitive unit (VNTR-MIRU) et restriction fragments length polymorphism (RFLP) - M.L. Boschirola (UZH).
- 2006-07 (6 mois) **Chargé de projets** - Institut Pasteur de Lille - Centre de typage moléculaire (IPL-CTM) - Lille, France :
En charge d'audits (Qualité et sécurité), supervision de projets de typage de *Listeria monocytogenes* par sérotypage, ribotypage et pulsotypage - J.P. Vincent (IPL-CTM).
- 2006 (6 mois) **Stagiaire en maîtrise** - Institut Pasteur - Centre de typage moléculaire (IPL-CTM) - Lille, France :
Organisation de la gestion informatique de souchiers de *Listeria monocytogenes* caractérisé par sérotypage, ribotypage et pulsotypage - J.P. Vincent (IPL-CTM).
- 2005 (2 mois) **Technicien de laboratoire** - Agence française de sécurité sanitaire des aliments (AFSSA) - Laboratoire de sécurité des aliment (LSAL) - Boulogne sur mer, France :
En charge d'analyses : total volatile acid base, Kjeldhal, microbiologie normalisée, random amplified polymorphism DNA (RAPD), polymerase chain reaction (PCR) - G. Bourdin (LSAL).
- 2005 (2 mois) **Stagiaire en licence** - Agence française de sécurité sanitaire des aliments : AFSSA - Laboratoire de sécurité des aliment (LSAL) - Boulogne sur mer, France :
Etude de la croissance de *Listeria monocytogenes* dans du saumon par tests de croissance à différentes températures - G. Bourdin (LSAL).
- 2002-04 (2 ans) **Technicien hygiéniste** - Institut Pasteur de Lille, service de microbiologie et hygiène alimentaire (IPL-SERMHA) - Caen, France :
En charge de prélèvements (eau et aliments), audits hazard analyses critical control point (HACCP), interprétation des résultats d'analyse microbiologiques, contrôle de la qualité des huiles - J.P. Vincent (IPL-SERMHA).
- 2002 (2 mois) **Stagiaire en diplôme universitaire et technologique** - Université de Caen - Équipe de recherche en physico-chimie et biotechnologie (ERPCB) - Caen, France :
Étude de la fermentation de *Pseudomonas putida* et développement du dosage de sucres et dérivés acides par high performance liquid chromatography (HPLC) - D. Corroler (ERPCB).

DIPLÔMES

- 2007-11 **Doctorat** en sciences et techniques - Université Paris-Est (UPE), École des ponts ParisTech (ENPC) (Champs sur Marne, France) - UMR MA-102 Laboratoire eau environnement et systèmes urbains (LEESU) :
Installation de laboratoires de microbiologie, veille bibliographique, développements analytiques et applications sur matrices réelles, rédaction de publications scientifiques et de rapports d'activités, communication en congrès internationaux et nationaux, enseignement, bactériologie, biologie moléculaire, bioinformatique, tests paramétriques et non-paramétriques.
- 2006-07 **Master** en santé publique (4^{ème}/86), mention qualité et gestion des risques sanitaires et environnementaux, spécialité sécurité alimentaire - Université de Lille (Loos, France) - Institut lillois d'ingénierie de la santé (ILIS) :
Évaluation des risques chimiques et microbiologiques, référentiels qualité, management.
- 2004-06 **Titre d'ingénieur maître, maîtrise** (6^{ème}/33) et **licence** (9^{ème}/34) en biotechnologies et bioindustries - Université du littoral côte d'opale (ULCO)- Institut universitaire et professionnelle en qualité des procédés agroalimentaires et halieutiques (IUP QPAH) :
Microbiologie et physico-chimie analytique, biologie moléculaire, fermentation, génie des procédés, analyse de composantes principales, modèles prédictifs.

2000-02 Diplôme universitaire et technologique en génie biologique (10^{ème}/28), option industries alimentaires et biologiques (IAB) - Université de Caen - Institut universitaire et technologique (IUT) en génie biologique (GB) de Caen (Caen, France) :

Chimie, microbiologie, biologie moléculaire, physiologie, histologie, statistique fondamentale.

FORMATIONS QUALIFIANTES

2011-14 Qualification aux fonctions de maître de conférences des universités (Paris, France) : Section 65 biologie cellulaire (N°14265227300, 05/02/2014), Section 67 biologie des populations et écologie (N°12267227300, 08/02/2012), Section 68 biologie des organismes (N°12268227300, 26/01/2012).

2013 Certificat de sélection du Québec (Montréal, Canada) : CNP 4122 Assistant d'enseignement et recherche postsecondaire (N°C0005780756, 23/04/2013).

AUTRES FORMATIONS

2018 Langage Python (Bordeaux, France) : CNRS Formation

(4 jours) Principes de programmation, codage itératif et objet, codage collaboratif, applications BioPython.

2018 Bases de données et langage SQL pour non-informaticiens (Paris, France) : ORSYS

(3 jours) Modèle de la base de données, principales et fonctions SQL, développements de requêtes SQL.

2017-2018 Statistique et langage R (Maisons-Alfort, France) : École de l'ADN

(6 jours) Méthodes avancées de traitement des données, représentations graphiques complexes.

2015-2019 Phylogénie moléculaire (Maisons-Alfort, France) : BioSciences and Co

(6 jours) Modèles d'évolution, maximum de vraisemblance, modèles Bayésiens, horloges moléculaires.

2017 Encadrement de doctorants (Maisons-Alfort, France) : Adoc Mètis

(3 jours) Analyse des risques, gestion des co-encadrements, comités de suivi, intégrité scientifique.

2017 Risques biologiques, flux et alertes (Maisons-Alfort, France) : ANSES

(2 jours) Nomenclature des risques en laboratoires, procédures d'urgence, extinction d'incendie, secours.

2006-2019 Bioinformatique : gestion de donnée (Bash, Python), base de données (SQL, XML), site web (Dreamweaver, Nvu), taxonomie (Bionumerics, BioEdit, MEGA, Best, MrBayes, RAXML, IQ-Tree), statistique (R, JMP, Statgraphics), conception d'amorces et sondes moléculaires (AlleID, Beacon designer), comparaison de génomes (SPAdes, MinH, MeDuSa, GapCloser, FastQC, Integrated Genome viewer, Burrows-Wheeler Aligner, Bowtie 2, Genome Analyser Tool Kit, Picard, Samtool, UnifiedGenotyper, HalotypeCaller, MycoHit).

CONCEPTION, RÉALISATION ET CONDUITE DE PROJETS DE RECHERCHE

2018-2019 Programmes de recherche (Maisons-Alfort, France) : Collaboration européenne (plusieurs pays) du projet 'risk and disease burden of antimicrobial resistance' (RaDAR, one health european joint project (EJP), grant #773830) - Participation aux conception et réalisation - Base de données exhaustives de plasmides complètement assemblés, description des populations de plasmides bactériens (Articles en cours).

2016-2019 Programmes de recherche (Maisons-Alfort, France) : Collaboration nationale (ANSES et INRA) concernant le projet 'Méthode automatique de typage bactérien par pangénomique et variants alléliques à partir de séquençage à haut débit' (grant #Typautobac) - Conception, soumission et réalisation - Études d'association et ontologie de gènes à l'échelle pangénomique, décryptage de l'adaptation à l'hôte de *Salmonella* sur la base de données de séquençage haut débit (2 Articles).

2015-2019 Programmes de recherche (Maisons-Alfort, France) : Collaboration européenne (plusieurs pays) des working package 4 et 7 financées par le projet 'collaborative management platform for detection and analyses of re-emerging and foodborne outbreaks in Europe' (COMPARE, Horizon 2020, grant #643476) - Participation aux conception et réalisation - Installation du séquençage haut débit et de la génomique bactérienne, développement d'un outil d'assemblage *de novo* (ARTwork), développement d'un outil d'identification de variants du coregénomique (iVARcall2), installation d'un réseau Linux d'analyse et stockage de données, formation Linux des collègues du laboratoire, application à *Salmonella*, *Listeria*, *Clostridium*, *Bacillus* et *Staphylococcus* (9 Articles).

2014-15 Subventions postdoctorales (Montréal, Canada) : Collaboration institutionnelle (plusieurs laboratoires de l'institut) financée par l'institut de recherche du centre de santé de l'université de McGill (RI MUHC) (9 000 \$ CAN net) et des fonds de recherche en santé Québec (FRQS) (30 000 \$ CAN net) dans le cadre du projet 'Tuberculosis in Nunavik' (grant #29836) - Participation aux conception, soumission et réalisation - Installations et développements d'analyses en génomique bactérienne, décryptage de l'évolution de *Mycobacterium tuberculosis* causant des épidémies dans le Nunavik au Canada (2 Articles).

- 2011-13 (29 mois)** **Subventions postdoctorales** (Montréal, Canada) : Collaboration institutionnelle (plusieurs laboratoires de l'institut) financée par l'institut de recherche du centre de santé de l'université de McGill (50 250 \$ CAN net) et des fonds de recherche en santé Québec (FRQS) (60 000 \$ CAN net) dans le cadre du projet 'Tuberculosis in Nunavik' (grant #26274) - Participation aux conception, soumission et réalisation - Installations et développements d'analyses en génomique bactérienne, développement d'outils pour diagnostiquer la maladie de Crohn, décryptage de la perte de virulence de *Mycobacterium tuberculosis* (2 Articles).
- 2007-11 (39 mois)** **Allocation doctorale et monitorat d'enseignement** (Marne-la-Vallée, France) : Collaboration nationale (LEESU, ENPC, Eau de Paris, APHP) financée par l'École des ponts ParisTech (65 139 € net) et l'université Paris-Est de Créteil (UPEC) (6 640 € net) dans le cadre du projet 'Sources des mycobactéries non-tuberculeuses dans les bassins versants' porté par le 'programme interdisciplinaire de recherche sur l'eau et l'environnement du bassin de la Seine' (PIREN-Seine), l'observatoire des polluants urbains (OPUR) et la ville de Paris (grant #CRECEP-STEAM) - Conception, soumission et réalisation - Développement de méthodes bactériologique et moléculaire, identification des sources de *Mycobacterium* dans l'environnement, comparaison de génomes pour améliorer les cibles moléculaires (5 Articles).
- 2007 (6 mois)** **Allocation de stage de master** (Maisons-Alfort, France) : Collaboration nationale (INRA et AFSSA) financée par le 'veterinary network of laboratories researching into improved diagnosis and epidemiology of mycobacterial diseases' (VENoMYC), l'INRA and l'AFSSA (grant #AIP P00297) - Réalisation - Description de la diversité génétique par typage moléculaire conventionnel, association de la spécificité d'hôte de *Mycobacterium avium* (1 Article).

COLLABORATIONS ET PARTENARIATS SCIENTIFIQUES

Période	Recherche de partenaires	Établissement de partenariats	Montage de conventions
2015-2019 (5 ans)	Partenariats pour développer des analyses en génomiques appliquées à <i>Salmonella</i> , <i>Listeria</i> , <i>Staphylococcus</i> , <i>Bacillus</i> et <i>Clostridium</i>	INRA-ANSES (M.Y. Mistou) INRA (M. Mariadassou) INRA (H. Chiapello) INRA (P. Velge)	Convention ANSES-INRA Typautobac (Méthode automatique de typage bactérien par pangéome et variants alléliques à partir de séquençage à haut débit)
2011-2015 (5 ans)	Partenariats pour étudier des épidémies humaines de <i>Mycobacterium tuberculosis</i> dans le Nunavik	RI MUHC (M.A. Behr) M.B. Reed (RI MUHC) VirginiaTech (J.O. Falkinham) Paris Sud (C. Sola)	Institut de recherche du centre de santé de l'université de McGill (RI MUHC) et fonds de recherche en santé Québec (FRQS)
2007-2011 (5 ans)	Partenariats pour étudier les sources des mycobactéries non-tuberculeuses dans les bassins versants	LEESU (F. Lucas, R. Moilleron) Eau de Paris (L. Moulin) CNR Mycobactéries (E. Cambau) AFSSA-UZB (M.L. Boschioli)	Programme interdisciplinaire de recherche sur l'eau et l'environnement du bassin de la Seine (PIREN-Seine) et observatoire des polluants urbains (OPUR)

CONTACTS DES RÉFÉRENCES

Dr Michel-Yves Mistou	INRAE-ANSES	Dr Emmanuelle Cambau	AP-HP
Dr Mahendra Mariadassou	INRAE	Dr Maria Laura Boschioli	ANSES-UZB
Dr Hélène Chiapello	INRAE	Dr Joseph O. Falkinham III	VirginiaTech
Dr Philippe Velge	INRAE	Dr Franck Biet	INRAE-UR1282
Dr Marcel A. Behr	RI MUHC	Dr Cécile Rousseau	IUT UPEC
Dr Michael .B. Reed	RI MUHC	Françoise Odelin	IUT UPEC
Dr Christophe Sola	Paris Sud	Dr Graziella Bourdin	ANSES-LSAL
Dr Régis Moilleron	LEESU	Jean-Pierre Vincent	IPL-CTM-SERMYA
Dr Françoise Lucas	LEESU	Dr David Corroler	ERPCB
Dr Laurent Moulin	Eau de paris		

ENCADREMENTS D'ÉTUDIANTS

Période	Niveau	Mois	Stagiaire	Nat.	Thématique (École - Université)	Taux (%)	Articles
2017-2019	PhD	36	Méryl Vila Nova	FR	Génomique bactérienne (École Doctorale ABIES)	100	1
2019	Master 2	6	Marie-Noel Mansour	LB	Phylogénie des <i>Salmonella</i> (Chimie alimentaire - St. Joseph de Beyrouth)	100	0
2018	Master 2	6	Kévin La	FR	Étude d'association large de génomes (Biologie Informatique - Paris Diderot)	100	1
2018	PhD	12	Lena Fritsch	DE	Bioinformatique et rédaction (École Doctorale ABIES)	20	1
2017-2018	PhD	48	Abakabir-Mahamat	TD	Bioinformatique appliquée à <i>Clostridium</i> (École Doctorale ABIES)	40	1
2017-2018	PhD	36	Yann Sévellec	FR	Bioinformatique appliquée à <i>Salmonella</i> (École Doctorale ABIES)	40	3
2016	Master 2	6	Méryl Vila Nova	FR	Algorithme de détection de variants fixés (Génomique Inform. Mathém. Santé - Paris Saclay)	100	1
2015	Licence 2	1	Audrey Artignan	FR	Littérature sur les taux d'évolution (Génétique Humaines - Univ. Collège de London)	100	0
2013-14	PhD	24	Robyn S. Lee	CA	Phylogénie et identification de SNPs (Épidémiologie - Université McGill)	20	2
2014	Master 1	6	Trevor Tarakjian	CA	Gestion des données haut-débit (Épidémiologie - Université McGill)	15	0
2013	Master 1	6	Andrei Dan	RO	Méthode de détection de <i>Leifsonia</i> (Épidémiologie - Université McGill)	40	0
2012	Licence 1	3	Jeremy Dabor	CA	Extraction d'ADN mycobactériens (Épidémiologie - Université McGill)	10	0
2011	Médecine	24	Louis Kreitmann	FR	Travaux en laboratoire (Médecine - Université de Lille)	10	1
2010	Master 1	6	Yacine Boudali	AG	Quantification de <i>Mycobacterium</i> (Sciences Génie Env. - Université Créteil)	80	0
2009	Master 2	6	Laetitia Betelli	FR	Quantification de <i>E. coli</i> et <i>Enterococcus</i> (Sciences Génie Env. - Université Créteil)	20	1

ANIMATION DE LA RECHERCHE

Co-organisation de séminaires

Période	Laboratoire - Institut	Sujet	Fréquence
2015-19	GAMeR - LSAL - ANSES	Débat autour de la bioinformatique et de la génomique (DEBUG) ouvert en interne et externe de l'agence	3-4 mois
2011-15	Behr Lab - RI MUHC	Séminaire de laboratoire sur le avancées des travaux de différentes unités de recherche de l'institut	2-3 mois
2007-11	LEESU - UMR MA102	Représentant des doctorants, en charge des séminaires de doctorant et rapports de conclusions des conseils de laboratoire	3-4 mois

Appartenance à des sociétés savantes

Société française de bioinformatique (SFBI).
 Société chimique américaine (ACS).
 Société française de microbiologie (SFM).
 Société française de mycobactériologie (Mycoclub).
 Association française d'écologie microbienne (AFEM).

Organisation de comités de thèse

Membres des deux comités de thèse de la doctorante Méryl Vila Nova (2017 et 2018) :

Pr. Hélène Chiapello (INRA, Unité de mathématiques et informatique appliquées, Toulouse).
 Arnaud Felten (ANSES, Génome analyse Modélisation risque, Maisons-Alfort).
 Valentin Loux (INRA, Plateforme de bioinformatique MIGALE, Jouy-en-Josas).
 Pr Philippe Velge (INRA, Infectiologie et santé publique, Nouzilly).
 Pr Daniel Wilson (John Radcliffe Hospital, Microbiology department, Oxford).

Membre de jurys de thèse de Doctorat

Année	Rôle	Étudiant en thèse	Université	Sujet de thèse
2019	Examineur	Damien Richard	Université de La Réunion	Adaptation génomique de <i>Xanthomonas citri</i> pv. <i>citri</i>
2018	Examineur	Franck Cerutti	Université de Toulouse III - Paul Sabatier	Dynamiques d'évolution de <i>Listeria monocytogenes</i>

Membre de comités de thèse de Doctorat

Période	Nombre	Étudiant en thèse	Université	Sujet de thèse
2017-2018	2	Maëllys Kevin	Paris-Est (UPE)	Diversité génomique de <i>Francisella tularensis</i> d'origine humaine et animale

Critique éditorial pour des journaux scientifiques

Microbial Genomics (1 éval. depuis 2019).
Frontiers in Microbiology (2 éval. depuis 2018).
Infection, Genetics and Evolution (1 éval. depuis 2015).
Environmental Science and Technology (1 éval. depuis 2014).
Diagnostic Microbiology and Infectious Disease (2 éval. depuis 2014).
PloS One (5 éval. depuis 2013).
Journal of Clinical Microbiology (12 éval. depuis 2012).
Applied and Environmental Microbiology (4 éval. depuis 2011).
Journal of Environmental Science and Health (2 éval. depuis 2011).
International Journal of Systematic and Evolutionary Microbiology (2 éval. depuis 2011).
Letters in Applied Microbiology (5 éval. depuis 2011).
Journal of Applied Microbiology (2 éval. depuis 2011).
Journal of Medical Microbiology, section Clinical Microbiology and Virology (2 éval. depuis 2010).

SÉJOURS DE COURTE DURÉE DANS DES LABORATOIRES DE RECHERCHE

2010 (3 mois) **Subvention de mobilité internationale** (Marne-la-Vallée, France) : Collaboration internationale (École des ponts ParisTech et institut polytechnique de virginie) financée par l'université Paris-Est (5 000 € net) dans le cadre du projet 'Sources des mycobactéries non-tuberculeuses dans les bassins versants' - Conception, soumission et réalisation - Transfert technologique d'une méthode de quantification moléculaire de *Mycobacterium* dans l'environnement (1 Chapitre de livre).

PARTICIPATION À DES CONGRÈS COMME AUDITEUR

Printemps des territoires : 6 juin 2019 à Paris (France).
 Genome microbial identifier #11 (GMI) : 18 au 20 mars 2018 à Genève (Suisse).
 Pathobiome 2018 : 18 au 20 mars 2018 à Ajaccio (France).
 Genome microbial identifier #10 (GMI) : 15 au 17 mai 2017 à Cabo San Lucas (Mexique).
 Genome microbial identifier #11 (GMI) and Food and Agriculture Organization : 24 au 27 avril 2016 à Rome (Italie).
 Pathobiome 2015 : 24 au 26 juin 2015 à Maisons-Alfort (France).

ENSEIGNEMENTS

Année	Niveau	Nature	Discipline	Volume (eq. TD)
2015	Ingénierie	TP-TD	Workflow bioinformatique	16 h
2010	Licence	TP-TD	Microbiologie médicale	64 h
2009	Master	Cours	Outils analytiques en épidémiologie	3 h
2009	Licence	TP	Microbiologie générale	64 h
2008	Licence	TP	Microbiologie générale	30 h

VALORISATION DE LA RECHERCHE

Publications internationales

- Douarre P.E., Mallet L., **Radomski N.**, Felten A. and M.Y. Mistou. Analysis of COMPASS, a new comprehensive plasmid database. 2020, under submission.
- Bonis M., Felten A., Payraud S., Dijoux A., Mallet L., **Radomski N.**, Mistou M.Y., Firmesse O., Hennekinne J.A. and S. Herbin. Characterization of *Bacillus thuringiensis* isolates associated to Foodborne Outbreaks in France from 2007 to 2017. 2020, under submission.
- Palma F., Brauge T., **Radomski N.**, Mallet L., Felten A., Mistou M.Y., Brisabois A., Guillier L. and G. Midelet-Bourdin. Dynamics of mobile genetic elements of *Listeria monocytogenes* persisting in ready-to-eat seafood processing plants in France. 2020, BMC Genomics, 21(1): 130, doi: 10.1186/s12864-020-6544-x (IF 2017 : 3.730).
- Vila Nova M, Durimel K., La K., Felten A., Bessières P., Mistou M.Y., Mariadassou M. and **N. Radomski**. Genetic and metabolic signatures of *Salmonella enterica* subsp. *enterica* associated with animal sources at the pangenomic scale. 2019, BMC Genomics, 20(1): 814, doi: 10.1186/s12864-019-6188-x (IF 2017 : 3.730).
- Radomski N.** †, Cadel-Six S. † († **equal contribution**), Cherchame E., Felten A., Barbet P., Palma F., Mallet L., Le Hello S., Weill F.X., Guillier L. and M.Y. Mistou. A simple and robust statistical method to define genetic relatedness of samples related to outbreaks at the genomic scale - Application to retrospective *Salmonella* foodborne outbreak investigations. 2019, Frontiers in Microbiology, 10(2413): 1-13, doi: 10.3389/fmicb.2019.02413 (IF 2017 : 4.019).
- Abakabir Mahamat A., **Radomski N.**, Delannoy S., Djellal S., Lenegrade M., Hadjab K., Fach P., Hennekinne J.A., Mistou M.Y. and O. Firmesse. Large-scale genomic analyses and toxinotyping of *Clostridium perfringens* implicated in foodborne outbreaks in France. 2019, Frontiers in Microbiology, 10(777): 1-14, doi: 10.3389/fmicb.2019.00777 (IF 2017 : 4.019).
- Sévellec Y., Felten A., **Radomski N.**, Granier S.A., Le Hello S., Petrovska L., Mistou M.Y. and S. Cadel-Six. Genetic diversity of *Salmonella* Derby from poultry sector in Europe. 2019, Pathogens, 8(2): piiE46, doi: 10.3390/pathogens8020046 (IF 2019 : 3.520).
- Fritsch L., A. Felten, F. Palma, J.F Mariet, **N. Radomski**, M.Y. Mistou, J.C. Augustin and L. Guillier. Insights from genome-wide approaches to identify variants associated to phenotypes at pan-genome scale: Application to *L. monocytogenes* ability to grow in cold conditions. 2018, International Journal of Food Microbiology, 291(16): 181-188, doi.org/10.1016/j.ijfoodmicro.2018.11.028 (IF 2017 : 3.451).
- Sévellec Y., M.L. Vignaud, S.A. Granier, R. Lailler, C. Feurer, S.L. Hello, M.Y. Mistou and S. Cadel-Six (acknowledgment L. Guillier and **N. Radomski**). Polyphyletic nature of *Salmonella enterica* serotype derby and lineage-specific host-association revealed by genome-wide analysis. 2018, Frontiers in Microbiology, 9(891): 1-13, doi.org/10.3389/fmicb.2018.00891 (IF 2017 : 4.019).
- Sévellec Y., S. A. Granier, **N. Radomski**, A. Felten, S. Le Hello, C. Feurer, M. Y. Mistou and S. Cadel-Six. Complete genome sequence of *Salmonella enterica* subsp. *enterica* serotype Derby, associated with the pork sector in France. 2018, Microbiology Resource Announcements, 7(12): e01027-18, doi: 10.1128/MRA.01027-18 (IF 2017 : 1.180).
- Felten A., M. Vila Nova, K. Durimel, L. Guillier, M.Y. Mistou and **N. Radomski**. First gene-ontology enrichment analysis based on bacterial coregenome variants: insights into adaptations of *Salmonella* serovars to mammalian- and avian-hosts. 2017, BMC Microbiology, 17(222): 1-20, doi.org/10.1186/s12866-017-1132-1 (IF 2015 : 2.960).
- Henri C., P. Leekitcharoenphon, H.A. Carleton, **N. Radomski**, R. S. Kaas, J. F. Mariet, A. Felten, F. M. Aarestrup, P. G. Smidt, S. Roussel, L. Guillier, M.Y. Mistou and R.S. Hendriksen. An assessment of different genomic approaches for inferring phylogeny of *Listeria monocytogenes*. 2017, Frontiers Microbiology, 8(2351): 1-13, doi.org/10.3389/fmicb.2017.02351 (IF 2016 : 4.076).
- Felten A., L. Guillier, **N. Radomski**, R. Lailler, M.Y. Mistou and S. Cadel-Six. Genome Target Evaluator (GTEvaluator): a workflow exploiting genome dataset to measure the sensitivity and specificity of genetic markers. 2017, PLoS ONE, 12(7): e0182082, doi: 10.1371/journal.pone.0182082 (IF 2015 : 4.411).
- Lee R.S. †, **N. Radomski** † († **equal contribution**), J.F. Proulx, I. Levade, B.J. Shapiro, F. McIntosh, H. Soualhia, D. Menzies and M.A. Behr. Population genomics of *Mycobacterium tuberculosis* in the Inuit. 2015, Proceedings of the National Academy of Sciences of the United States of America, PNAS, 112(44): 13609-13614, doi: 10.1073/pnas.1507071112 (IF 2014 : 9.674).
- Wang J., F. McIntosh, **N. Radomski**, K. Dewar, R. Simeone, J. Enninga, R. Brosch, E.P. Rocha, F.J. Veyrier and M.A. Behr. Insights on the emergence of *Mycobacterium tuberculosis* from the analysis of *Mycobacterium kansasii*. 2015, Genome Biology and Evolution, 7(3): 856-870, doi: 10.1093/gbe/evv035 (IF 2013 : 4.532).
- Lee R.S. †, **N. Radomski** † († **equal contribution**), J.F. Proulx, J. Manry, F. McIntosh, F. Desjardins, H. Soualhia, P. Domenech, M.B. Reed, D. Menzies and M.A. Behr. Re-emergence and amplification of tuberculosis in the Canadian Arctic. 2015, Journal of Infectious Diseases, pii(jiv011): 1-10, doi: 10.1093/infdis/jiv011 (IF 2013 : 5.778).

Domenech P., A. Rog, J. Moolji, **N. Radomski**, A. Fallow, L. Leon-Solis, J. Bowes, M.A. Behr and M.B. Reed. The origins of a 350-kilobase genomic duplication in *Mycobacterium tuberculosis* and its impact on virulence. 2014, *Infection and Immunity*, 82(7): 2902–2912, doi: 10.1128/IAI.01791-14 (IF 2012 : 4.074).

Radomski N., A. Roguet, F.S. Lucas, F. J. Veyrier, E. Cambau, H. Accrombessi, R. Moilleron, M.A. Behr and L. Moulin. *atpE* gene as a new useful specific molecular target to quantify *Mycobacterium* in environmental samples. 2013, *BMC Microbiology*, 13(277): 1-12, doi: 10.1186/1471-2180-13-277 (IF 2012 : 3.104).

Radomski N., L. Kreitmann, F. McIntosh and M.A. Behr. The critical role of DNA extraction for detection of mycobacteria in tissues. 2013, *PLoS ONE*, 8(10): e78749, doi: 10.1371/journal.pone.0078749 (IF 2012 : 3.703).

Radomski N., L. Betelli, R. Moilleron, S. Haenn, L. Moulin, E. Cambau, V. Rocher, A. Gonçalves and F.S. Lucas. *Mycobacterium* behavior in wastewater treatment plant, a bacterial model distinct from *Escherichia coli* and enterococci. 2011, *Environmental Science & Technology*, 45(12), 5380–5386, doi: 10.1021/es104084c (IF 2009 : 4.630).

Radomski N., F.S. Lucas, R. Moilleron, E. Cambau, S. Haenn and L. Moulin. Development of a real-time qPCR method for detection and enumeration of *Mycobacterium* spp. in surface water. 2010, *Applied and Environmental Microbiology*, 76(21), 7348-7351, doi: 10.1128/AEM.00942-10 (IF 2009 : 3.686).

Radomski N., E. Cambau, L. Moulin, S. Haenn, R. Moilleron and F.S. Lucas. Comparison of culture methods for isolation of nontuberculous mycobacteria from surface waters. 2010, *Applied and Environmental Microbiology*, 76(11), 3514-3520, doi: 10.1128/AEM.02659-09 (IF 2009 : 3.686).

Radomski N., V.C. Thibault, C. Karoui, K. de Cruz, T. Cochard, C. Gutiérrez, P. Supply, F. Biet and M.L. Boschirolì. Genotypic diversity of *Mycobacterium avium* subspecies from human and animal origins, studied by MIRU-VNTR and IS1311 RFLP typing methods. 2010, *Journal of Clinical Microbiology*, 48(4), 1026-1034, doi: 10.1128/JCM.01869-09 (IF 2009 : 4.162).

Graziella Midelet-Bourdin, Stéphanie Copin, Guylaine Leleu and Pierre Malle (acknowledgment Céline Sart, Mylène Gobert, **Nicolas Radomski** and Rémi Cappelaere). Determination of *Listeria monocytogenes* growth potential on new fresh salmon preparations. 2010, *Food Control*, 21(10), 1415-1418, <https://doi.org/10.1016/j.foodcont.2010.03.009> (IF 2010 : 2.812).

Publications nationales

Radomski N. Sources des mycobactéries non-tuberculeuses dans les bassins versants. Thèse de doctorat en sciences et techniques de l'environnement. Soutenance : 28 février 2011 sous la direction de Régis Moilleron - Université Paris-Est (UPE), doi: 10.13140/2.1.3436.6401 (<http://www.theses.fr/15836869X>).

Communications orales

Douarre P.E., Mallet L., **Radomski N.**, Felten A. et M.Y. Mistou. Global analysis of beta-lactams resistant plasmids (EN): 15^e congrès national de la société française de microbiologie (SFM): 30 septembre au 2 octobre 2019 à Paris (France).

Vila Nova M., Durimel K., La K., Felten A., Bessières F., Mistou M.Y., Mariadassou M. et **N. Radomski**. Adaptation to animal sources of *Salmonella enterica* subsp. *enterica* at the pangenomic scale (EN): Food Safety WGS: 26 au 28 mars 2019 à Paris (France) et IWBBIO: 9 au 10 mai 2019 à Grenade (Espagne) et JOBIM : 2 au 5 juillet 2019 à Nantes (France).

Radomski N., Cadel-Six S., Cherchame E., Felten A., Barbet P., Vignaud M.L., Mallet L., Le Hello S., L. Guillier et M.Y. Mistou. A simple and robust statistical method to define genetic relatedness of samples related to outbreaks at the genomic scale - Application to retrospective *Salmonella* foodborne outbreak investigations (EN): COMPARE general meeting (Work packages WP4 et WP7): 27 février au 01 mars 2019 à Lyngby (Danemark).

Radomski N., M. Vila Nova, K. Durimel, L. Guillier, M.Y. Mistou et **N. Radomski**. Metabolic processes mainly impacted by fixed coregenome variants during adaptations of *Salmonella* serovars to mammalian- and avian-hosts (EN): COMPARE meeting (Work packages WP4 et WP7): 22 au 23 mai 2018 à l'institut Robert Koch de Wernigerode (Allemagne).

Radomski N., S. Cadel-Six, E. Cherchame, S.L. Hello, L. Guillier et M.Y. Mistou. Foodborne outbreak (FBO) investigations: Is there a faster method than growing a tree? (EN): COMPARE general meeting (Work packages WP4 et WP7) : 28 février au 2 mars 2018 à Lyngby (Danemark).

Radomski N. Développements Bioinformatiques : Génomique bactérienne à l'ANSES et premiers résultats COMPARE sur *Salmonella* Typhimurium en France (EN) : Réunion générale COMPARE (Work packages WP4 et WP7) : 1 au 2 mars 2017 à Rotterdam (Pays-Bas).

***Radomski N.** R.S. Lee et M.A. Behr. Inférence de l'évolution *in natura* de *Mycobacterium tuberculosis* pendant un siècle de transmissions (EN) : Phylogénie Intégrative du 1^{er} Congrès "Réseau de Systématique / Muséum national d'Histoire naturelle" (R-Syst / MNHN) : 12 au 14 septembre 2015 in Versailles-Grignon (France).

Radomski N., R.S. Lee et M.A. Behr. Évolution de *Mycobacterium tuberculosis* pendant un siècle de transmissions (EN) : DEBUT-5 de l'agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail : 10 septembre 2015 à Maisons-Alfort (France).

‡**Radomski N.**, R.S. Lee, J.F. Proulx, F. McIntosh, H. Soualhine, P. Domenech, M. Reed, D. Menzies, et M. A. Behr. Diversité, évolution et distribution de *Mycobacterium tuberculosis* au Nunavik (EN) : Centre international sur la tuberculose de McGill : 18 juin 2014 à Montréal (Canada).

‡**Radomski N.** et M.A. Behr. Identification à l'aveugle de SNPs par séquençage haut débit de génomes : Description d'épidémie(s) de tuberculose chez les Inuits (EN) : Centre international sur la tuberculose de McGill : 15 janvier 2014 à Montréal (Canada).

‡**Radomski N.** et M.A. Behr. Le rôle critique de l'extraction d'ADN pour la détection des mycobactéries dans les tissus (EN) : 1^{ère} Journée annuelle des jeunes chercheurs du centre international sur la tuberculose de McGill : 19 juin 2013 à Montréal (Canada).

Radomski N. Écologie, Physiologie, Évolution, Épidémiologie et interactions hôte-bactérie d'Actinobacteria pathogènes majeures: Les mycobactéries (FR) : Séminaire de laboratoire à UMR 7267 Écologie et biologie des interactions, Équipe microbiologie de l'eau de l'université de Poitiers : 14 mai 2013 à Poitiers (France).

‡**Radomski N.** et M.A. Behr. Quantification de *Mycobacterium avium* spp. *paratuberculosis* (MAP) dans des tissus : extraction d'ADN, PCR en temps-réel versus méthode de culture (EN) : 5^{ème} réunion Canadienne annuelle sur MAP : 11 au 13 octobre 2012 à Banff (Canada).

Radomski N. Écologie, physiologie, évolution, et interactions hôtes-pathogène d'Actinobactéries majeures : Les mycobactéries : Séminaire de laboratoire à UMR 5557 Écologie microbienne de l'université Claude Bernard (FR) : 22 mai 2012 à Lyon (France).

***Radomski N.**, L. Betelli, R. Moillon, S. Haenn, L. Moulin, E. Cambau, A. Pruden, J.O. Falkinham III, V. Rocher, A. Gonçalves, et F.S. Lucas. Sources et comportements des mycobactéries non-tuberculeuses dans les bassins versants (FR) : Congrès du programme Interdisciplinaire de recherche sur l'eau de la Seine : 7 au 9 février 2011 à Paris (France).

*Mouchel J.M., J. Passerat, K. Ouattara, P. Servais, S. Ayrault, C. Priadi-Rianti, C. Gourlay, E. Uher, [...], J. Eurin, F. Alliot, A. Desportes, C. Bourges, M. Chevreuil, G. Varrault, Y. Louis, C. Lorgeoux, et **N. Radomski**. La rivière à la traversée de la ville : le temps de pluie dans l'agglomération parisienne (FR) : Congrès du programme interdisciplinaire de recherche sur l'eau de la Seine : 7 au 9 février 2011 à Paris (France).

***Radomski N.**, S. Haenn, R. Moillon, F.S. Lucas, E. Cambau, et L. Moulin. Développement de méthodes de quantification des mycobactéries non-tuberculeuses dans l'eau de la Seine par bactériologie et biologie moléculaire (FR) : Congrès du programme interdisciplinaire de recherche sur l'eau de la Seine : 27 au 28 janvier 2010 à Paris (France).

*‡**Radomski N.**, V. Thibault, C. Karoui, K. De Cruz, T. Cochard, C. Gutiérrez, P. Supply, F. Biet, et M.L. Boschirolu. Diversité génétique de souches de *Mycobacterium avium* spp. d'origines humaine et animale par caractérisation MIRU-VNTR et IS1311 RFLP (EN, FR) : 30^{ème} congrès annuel de la société européenne de mycobactériologie : 5 au 8 Juillet 2009 à Porto (Portugal) et 11^{ème} Journée de mycobactériologie de langue française : 15 au 16 octobre 2009 à Bandol (France).

Radomski N., F.S. Lucas, E. Cambau, L. Moulin, S. Haenn, et R. Moillon. Etude des sources des mycobactéries atypiques dans le bassin de la Seine à Paris (FR) : Congrès de l'association francophone d'écologie microbienne : Méthodes de culture des bactéries non-cultivables : 16 au 17 Juin 2008 à Banyuls (France).

*Gourlay-Francé C., S. Ayrault, M. Chevreuil, A. Da Silva, J. Eurin, P. Labadie, F. Lucas, [...], **N. Radomski**, V. Rocher, P. Servais, E. Uher, G. Varrault et J.M. Mouchel. Contamination microbiologique et chimique en Seine à la suite d'un rejet urbain de temps de pluie (FR) : Congrès du programme interdisciplinaire de recherche sur l'eau de la Seine : 5 au 6 février 2009 à Paris (France).

* Actes de congrès (n = 8). ‡ Invitations dans des congrès internationaux (n = 9). # Prix de congrès (n = 1).

Communications affichées

Palma F., Brauge T., **Radomski N.**, M.Y. Mistou, A. Brisabois, L. Guillier et Midelet-Bourdin G. Éléments génétiques mobiles à travers les complexes clonaux de *Listeria monocytogenes* contaminant les industries de transformation des produits de la mer en France (EN) : Réunion générale COMPARE (Work packages WP4 et WP7) : 27 février au 01 mars 2019 à Lyngby (Danemark).

Vila Nova M., K. La, A. Felten, K. Durimel, P. Bessieres, M.Y. Mistou, M. Mariadassou et **N. Radomski**. Adaptations aux hôtes de serovars de *Salmonella enterica* spp. *enterica* décryptées par la première étude d'association génomique implémentant les gènes accessoires et variants du coregénom (EN) : 16 au 18 mai 2018 à Genève (Suisse).

*Felten A., M. Vila Nova, K. Durimel, L. Guillier, M.Y. Mistou et **N. Radomski**. Première analyse d'enrichissement d'ontologie de gènes basée sur des variants du coregénom bactérien : A travers l'adaptation des mammifères et

aviaires aux sérovars de *Salmonella* (EN): Journées ouvertes en biologie, informatique et mathématiques (JOBIM) : 3-6 juin 2017 à Lille (France) et Bioinformatiques des algorithmes aux applications (BiATA) : 1 au 5 août 2017 à St. Petersburg (Russie).

* ‡#**Radomski N.**, R.S. Lee, F. McIntosh, P. Domenech, M. Reed, D. Menzies et M.A. Behr. Évolution de *Mycobacterium tuberculosis* dans la *natura*: Cas unique du Nunavik, Québec (EN) : Initiatives pour la santé mondiale, Club de la faculté McGill : 3 novembre 2014 à Montréal (Canada).

Felten A., L. Guillier, **N. Radomski**, M.Y. Mistou, R. Lailler, et S. Cadel-Six. Genome Target Selector : Estimation de la spécificité et sensibilité de cibles moléculaires basée sur des séquences de génomes bactériens (EN) : 11^{ème} Réunion international sur les marqueurs épidémiologiques microbiens (IMMEM XI). Navigation des génomes microbiens. Une conférence ESCMID - ESGEM : 9 au 12 mars 2016 à Estoril (Portugal).

* ‡#Lee R.S., **N. Radomski**, J.F. Proulx, J. Manry, F. McIntosh, F. Desjardins, H. Soualhiné, P. Domenech, M. Reed, D. Menzies et M.A. Behr. Réémergence et amplification de la tuberculose dans l'Arctique Canadien (EN) : Initiatives pour la Santé Mondiale de la faculté McGill : 3 novembre 2014 à Montréal (Canada).

‡Lee R.S., **N. Radomski**, F. McIntoch, F. Desjardins, H. Soualhiné, J.F. Proulx, D. Menzies et M.B. Behr. Le séquençage complet de génomes révèle des épidémies multiples de TB dans un village Canadien (EN) : La 18^{ème} conférence de l'union des régions Nord-Américaines « Plus forts ensemble pour stopper TB » du laboratoire à la clinique : 27 février au 1^{er} mars 2014 à Boston (USA).

* ‡Lucas F.S., **N. Radomski**, A. Roguet, E. Cambau, R. Moilleron, M.A. Behr, et L. Moulin. Les méthodes de détection pour l'étude des mycobactéries non-tuberculeuses dans les systèmes aquatiques sous pression du changement global (EN) : Congrès EMBO SAME13 : 8 au 13 septembre 2013 à Stresa (Italie).

Moulin L., S. Haenn, S. Dubrou, J.L. Gaillard, **N. Radomski**, E. Cambau, B. Welté, et M. Joyeux. Étude de la diversité des mycobactéries atypiques du système de distribution de l'eau (FR) : 27^{ème} congrès annuel sur l'eau de Eau de Paris : 24 au 28 février 2012 à Paris (France).

* ‡**Radomski N.**, E. Cambau, L. Moulin, S. Haenn, R. Moilleron, et F.S. Lucas. Détection des mycobactéries non-tuberculeuses dans l'eau de surface : Comparaison des méthodes de culture (EN, FR) : 30^{ème} congrès annuel de la société Européenne de mycobactériologie : 5 au 8 juillet 2009 à Porto (Portugal) et 11^{ème} Journée de mycobactériologie de langue française : 15 au 16 octobre 2009 à Bandol (France).

* ‡Lucas F.S., **N. Radomski**, E. Cambau, L. Moulin, S. Haenn, et R. Moilleron. Quantification des mycobactéries non-tuberculeuses dans l'eau de surface (FR) : Développement de cibles moléculaires : congrès annuel de la société Européenne de mycobactériologie : 5 au 8 juillet 2009 à Porto (Portugal).

***Radomski N.**, S. Haenn, R. Moilleron, F.S. Lucas, E. Cambau, et L. Moulin. Développement d'une méthode de biologie moléculaire pour quantifier les mycobactéries non-tuberculeuses dans l'eau de surface (FR) : Congrès du Programme Interdisciplinaire de Recherche sur l'Eau de la Seine : 5 au 6 février 2009 à Paris (France).

* Actes de congrès (n = 8). ‡ Invitations dans des congrès internationaux (n = 6). # Prix de congrès (n = 2).

Chapitre d'ouvrage

Radomski N., R. Moilleron, F.S. Lucas and J.O. Falkinham III. Challenges in environmental monitoring of pathogens: Case study in *Mycobacterium avium* (2011) Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology (In.), A. Méndez-Vilas (Ed.), (2), 1551-1561.

Brevets

Cadel-Six S., A. Felten, P. Fach, L. Guillier, **N. Radomski**, M.Y. Mistou et R. Lailler. ANSES - Laboratoire de sécurité des aliments (Unité SEL / Mission GAMeR / Plateforme IdentityPath). Algorithme de sélection de kmers sensibles et spécifique à l'échelle génomique et vérification des cibles moléculaire par PCR haut débit pour identifier les sérovars de *Salmonella*. 2018 (en cours d'examen).

Récompenses

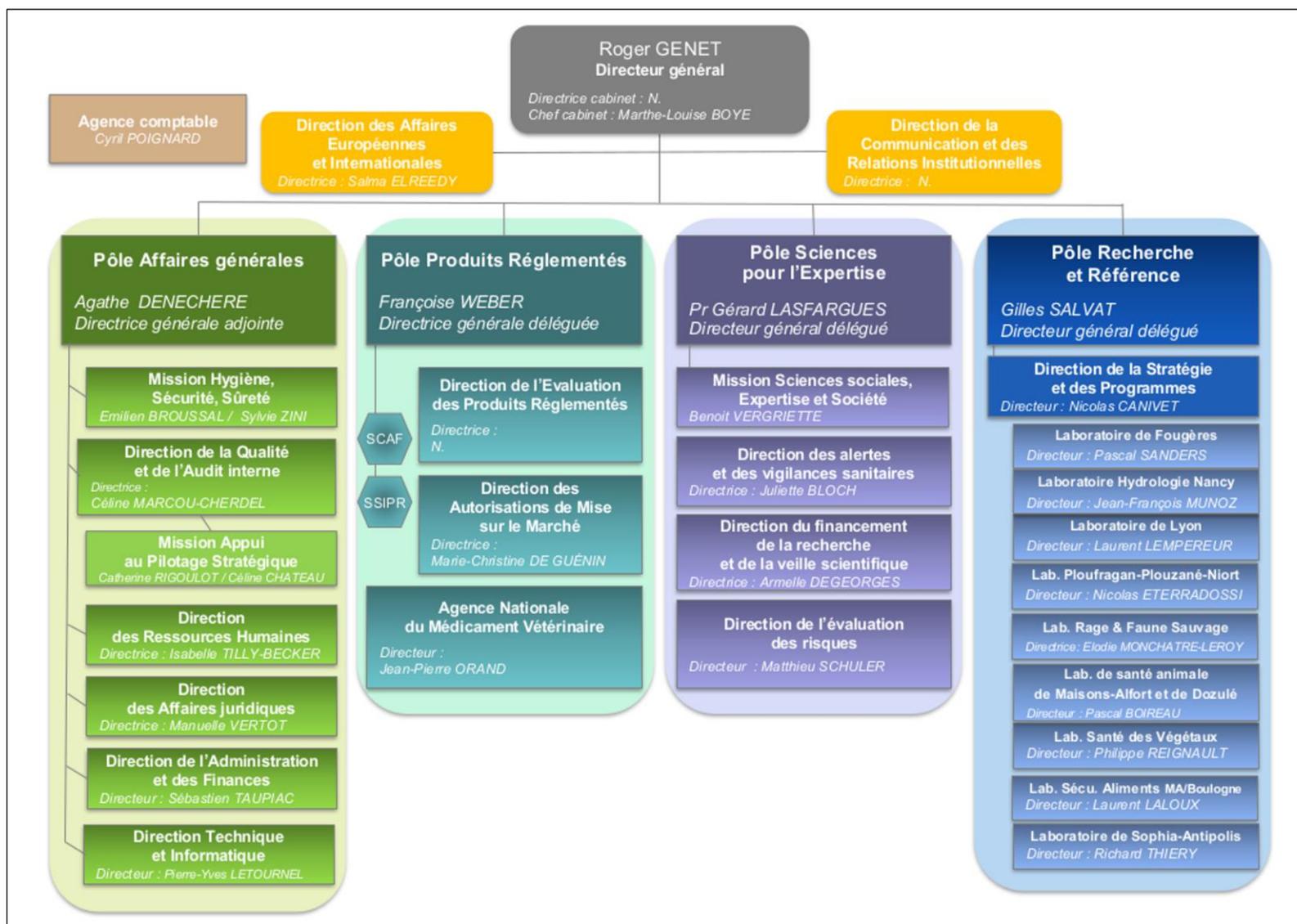
2014 **Prix de congrès** (Montréal, Canada) : Centre international sur la tuberculose de McGill :(Oral, 18/06/2014) et initiatives pour la santé mondiale (Poster, 03/11/2014).

2012 **Thèse de doctorat classée parmi les 48 meilleurs de 2011** (Paris, France) : Institut des sciences et technologies (ParisTech), Université Paris-Est (UPE).

Annexe 2 : Liste des établissements pouvant établir des relations conventionnelles avec l'agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (ANSES)

- Agence de l'environnement et de la maîtrise de l'énergie
- Agence nationale de sécurité des médicaments et des produits de santé
- Agence nationale pour l'amélioration des conditions de travail
- Bureau de recherches géologiques et minières
- Caisse centrale de la mutualité sociale agricole
- Centre international de recherche agronomique pour le développement
- Centre national de la recherche scientifique
- Centre scientifique et technique du bâtiment
- Institut national de recherche en sciences et technologies pour l'environnement et l'agriculture
- Commissariat à l'énergie atomique et aux énergies alternatives
- École des hautes études en santé publique
- École nationale vétérinaire de Maisons-Alfort
- École nationale vétérinaire de Toulouse
- École nationale vétérinaire, agroalimentaire et de l'alimentation, Nantes-Atlantique
- Institut d'enseignement supérieur et de recherche en alimentation, santé animale, sciences agronomiques et de l'environnement
- Institut de recherche pour le développement
- Institut des sciences et industries du vivant et de l'environnement
- Institut français de recherche pour l'exploitation de la mer
- Institut français des sciences et technologies des transports, de l'aménagement et des réseaux
- Institut national du cancer
- Institut national de la recherche agronomique
- Institut national de recherche et de sécurité pour la prévention des accidents du travail et des maladies professionnelles
- Institut national de la santé et de la recherche médicale
- Institut national de l'environnement industriel et des risques
- Institut de radioprotection et de sûreté nucléaire
- Institut Pasteur
- Laboratoire national de métrologie et d'essais
- Office national de l'eau et des milieux aquatiques
- Santé publique France

Annexe 3 : Organigramme hiérarchique de l'agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (ANSES) au 1^{er} avril 2019



Annexe 4 : Plaquette de présentation des trois plateformes d'épidémiosurveillance et de leurs acteurs sous l'égide du ministère des solidarités et de la santé et du ministère de l'agriculture et de l'alimentation



La plateforme d'épidémiosurveillance en santé animale couvre tout danger sanitaire ayant ou pouvant avoir un impact sur la santé animale et/ou la santé publique (zoonoses), et pour lequel une surveillance est souhaitable ou requise chez les animaux, à l'échelon de tout ou partie du territoire national. Cela concerne les animaux d'élevage, de compagnie ou de sport, ainsi que la faune sauvage.

Depuis sa création fin 2011, la Plateforme ESA a été mobilisée sur les épizooties qui ont touché le pays (par ex. maladie de Schmallenberg, fièvre catarrhale ovine, influenza aviaire hautement pathogène), les dangers sanitaires enzootiques (par ex. tuberculose bovine) ou qui le menacent (par ex. peste porcine africaine). En 2018, une vingtaine de dangers sanitaires faisaient partie de programme d'activité de la Plateforme ESA dans différentes filières de production (ruminants, porcs, volailles, abeilles). Une veille sanitaire en santé animale est opérationnelle depuis 2013 et met en ligne un bulletin hebdomadaire de veille.

Membres des trois plateformes d'épidémiosurveillance



The diagram illustrates the three platforms and their members:

- SCA Platform (purple):** Santé publique France, CGAD, OQUALIM, DGCCRF, fcd, ania, COP, ADIVA.
- ESV Platform (green):** acta, ANRS, ANSES, ANSES Santé Animale, ANSES Santé Publique, ANSES Santé Environnement, ANSES Santé Sécurité.
- ESA Platform (orange):** INRA, ANSES Santé Animale, ANSES Santé Publique, ANSES Santé Environnement, ANSES Santé Sécurité, SNQV, CIRAD, GDS.



Trois plateformes d'épidémiosurveillance

« *Une seule santé* »

Trois plateformes d'épidémiosurveillance sont désormais en place dans les domaines de la santé animale, de la santé végétale et de la surveillance de la chaîne alimentaire.

Ces plateformes rassemblent, dans un mode d'organisation public-privé à gouvernance partagée, l'ensemble des acteurs impliqués dans la surveillance des dangers sanitaires : Etat et représentants de ses délégués pour cette mission, organismes d'appui scientifique, instituts techniques agricoles, représentants des agriculteurs et des professionnels des filières de production, de transformation, de distribution et de restauration. Le partenariat public-privé permet d'optimiser les actions et les coûts de la surveillance, par un partage de ressources, de compétences et d'outils dans un objectif commun de protection de la santé animale, végétale et humaine.

Elles ont pour objectif de contribuer à la prévention et à la maîtrise des risques sanitaires, en veillant à ce que la surveillance épidémiologique dans les trois domaines soit la plus efficiente possible, de la production primaire au consommateur.

La surveillance des dangers sanitaires, microbiologiques et chimiques, et la mise en œuvre de dispositifs de détection rapide d'urgences doivent être menées de manière harmonisée et coordonnée entre les trois domaines, tant les similarités sont importantes et les interactions nombreuses, directement, via l'environnement et via les aliments. La nécessaire mise en commun de connaissances et de compétences contribue ainsi à construire une stratégie nationale pour l'épidémiologie en phase avec le concept *One Health* (« Une seule santé »).

A cette fin, une organisation inter-plateformes a été instaurée afin de favoriser les synergies et la continuité dans les collaborations entre plateformes.

Les objectifs de cette organisation sont :

Développer une culture collective commune aux différents domaines, sur des notions et concepts transversaux par nature (concept *One Health*, qualité des données, etc.).

Identifier les problématiques sanitaires nécessitant la mise en œuvre de dispositifs de surveillance intégrée entre les différents domaines.

Mettre en œuvre les collaborations nécessaires sur ces problématiques transversales : évaluation des dispositifs de surveillance, développements d'indicateurs sanitaires de l'amont à l'aval, production de tableaux de bord partagés de la situation sanitaire, etc.

Mutualiser des développements technologiques, de procéder à des partages d'expérience, à des transferts de savoir-faire, de compétences, ou de technologies.

Une cellule de coordination inter-plateformes, rassemblant les équipes de coordination des trois plateformes, a été constituée pour décliner et mettre en œuvre ces différents objectifs.

 Le champ d'action de la plateforme d'épidémiologie et de surveillance en santé végétale couvre potentiellement tout danger sanitaire ou phénomène phytosanitaire ayant ou pouvant avoir un impact sur l'état sanitaire des végétaux, et les effets non intentionnels des pratiques agricoles sur l'environnement. Elle mobilise les acteurs de la surveillance en appui aux responsables de dispositifs de surveillance sur des sujets méthodologiques (structuration des plans de surveillance officielle suite au nouveau règlement européen sur la santé des végétaux), l'amélioration des bilans sanitaires annuels et des thématiques sanitaires d'intérêt prioritaire pour la surveillance (en 2019 : *Xylella fastidiosa*, nématode du pin, dépérissement du vignoble). Des groupes de travail techniques et multi-partenaires sont constitués sur ces différents sujets. Une équipe opérationnelle mixte Inra-Anses apporte un appui scientifique et technique à ces travaux et se chargera d'activités transversales (site web et travaux exploratoires pour une veille sanitaire internationale par exemple).

 La plateforme de surveillance de la chaîne alimentaire inclut tous les stades de la chaîne alimentaire et tous les dangers (biologiques, chimiques ou physiques) susceptibles d'être présents dans les denrées d'origine animale et végétale, et pouvant représenter un risque pour l'Homme. L'alimentation animale est également incluse dans le périmètre de la plateforme. Dans une approche fondée sur la collaboration, les interactions et le consensus, la Plateforme SCA mobilise des acteurs de la production, de la transformation et de la distribution, ainsi que de la santé humaine. Des groupes de travail sont actuellement constitués pour œuvrer à l'optimisation de la surveillance des *Salmonella* en filière fromage au lait cru, ainsi qu'à l'amélioration des dispositifs de surveillance des *Salmonella* tout au long de la chaîne (GT Ondes), et notamment à l'abattoir avec la mutualisation des résultats d'autocontrôles. La Plateforme SCA ouvrira prochainement des travaux sur la veille sanitaire internationale, sur la surveillance des dangers chimiques et dangers microbiologiques autres que *Salmonella*, ainsi que sur la qualité des données de surveillance en transversalité avec les deux autres plateformes.

Annexe 5 : Mandats de laboratoire national de référence (LNR : cercle blanc) et laboratoire de référence de l'union européenne (LRUE : cercle noir) des sites Maisons-Alfort (MA) et Boulogne-sur-Mer (BsM) du laboratoire de sécurité des aliments (LSAL) organisé en départements, unités et missions transversales (T)

Départements	Unités ou Missions (lieu)	Mandats de LNR (○) et LRUE (●) dans les équipes											
		<i>Salmonella</i>	Staphylocoques à coagulase positive	<i>Listeria monocytogenes</i>	<i>Vibrio</i>	Botulisme aviaire	Anisakidae	Antibiorésistance	Pesticides (alimentaires)	Pesticides (monorésidus)	Métaux lourds	Biotoxines marines	Histamine
Contaminants microbiologiques des aliments	Unité <i>Salmonella</i> , <i>E. Coli</i> , <i>Listeria</i> (MA)	○		○									
	Unité Staphylocoques, <i>Bacillus</i> , <i>Clostridies</i> (MA)		○	●		○							
	Unité Laboratoire central des services vétérinaires (MA et BsM) T	○	○	○	○	○	○	○					
	Unité Virus entériques (MA)												
	Mission identification et typage des agents pathogènes (MA) T	○	○	○	○	○	○	○					
	Mission antibiorésistance (MA) T	○	○	○	○	○	○	○					
	Mission génome analyse modélisation risque (MA) T	○	○	○	○	○	○	○					
Produits de la pêche et de l'aquaculture	Unité bactériologie et parasitologie des produits de la pêche et de l'aquaculture (BsM)				○		○	○					
	Unité physico-chimie des produits de la pêche et de l'aquaculture (BsM)												
	Unité biochimie des produits aquatiques (BsM)												
Contaminants chimiques des aliments	Unité pesticides et biotoxines marines (MA)							○	○		○	○	
	Unité éléments traces métalliques et minéraux (MA)									○			

Annexe 6 : Note d'organisation de la mission génome analyse modélisation risque (GAMeR) - LSA-NO-0023 - LSA-MM-0001 - Manuel Qualité A

		<h3>Note d'organisation</h3>	
<h4>Organisation fonctionnelle de la mission « Genome Analyse Modélisation Et Risque »</h4>			
<i>Entité émettrice</i>	Laboratoire de Sécurité des Aliments Sites de Maisons-Alfort et de Boulogne-sur-Mer		APPROBATION Date : 12/02/2016 11:53:48 Laurent LALOUX
<i>Nature de la modification</i>	Réorganisation de la mission Mod AQR, de ses activités et de sa dénomination.		
<i>Diffusion pour application</i>	Personnel du laboratoire Direction des laboratoires (DL)		

Sommaire

1.	OBJECTIFS DE LA MISSION	1
2.	ORGANISATION FONCTIONNELLE	2
3.	MODE DE FONCTIONNEMENT	2
4.	ORGANIGRAMME FONCTIONNEL	2

1. Objectifs de la mission

La mission rassemble des compétences dans la mise en œuvre des méthodes bio-informatiques et statistiques d'analyse des données omiques associée à une expertise dans le domaine de la modélisation du risque microbiologique dans les aliments. Ses objectifs sont rassemblés ci-dessous sans ordre de priorité.

1 - Accompagner les composantes du laboratoire dans leurs projets de détection, typage et caractérisation des microorganismes pathogènes alimentaires fondées sur l'analyse des génomes.

2 - Faire évoluer les modèles utiles à l'appréciation de l'exposition des consommateurs aux dangers microbiologique, à l'attribution des sources, à la relation dose-réponse. Intégrer les données omiques.

Anses LSAI - LSA n	NOTE D'ORGANISATION GAMeR LSA-NO-0023 (LSA-MM-0001 – Manuel Qualité A)	Révision	01	1/2
		Date	12/02/2016	

Seuls la version informatique et les exemplaires en diffusion contrôlée font foi.

3 - Développer une recherche de haut niveau dans le domaine de la génomique/transcriptomique des bactéries pathogènes alimentaires, à travers le support des projets portés dans les équipes du laboratoire et en définissant son propre programme de recherche.

4 - Faire diffuser l'expertise en analyse des données omiques au sein du laboratoire par des actions d'animation et de formation. La mission pourra également contribuer aux actions au niveau national initiée par la plate-forme nationale de séquençage. Assurer la veille technologique dans ses domaines de compétence.

2. Organisation fonctionnelle

La mission est intégrée au département des contaminants microbiologiques des aliments et est dirigée par le chef de ce département. Elle est sollicitée par l'ensemble des composantes du laboratoire (départements, unités, missions, plate-forme) à travers une fiche projet.

La mission maintient des liens étroits avec la plateforme nationale de séquençage de Ploufragan afin d'optimiser les ressources humaines et techniques en bio-informatique à l'agence, mais également pour l'utilité des échanges entre experts. Elle s'attachera à participer activement à un réseau Anses si d'autres bio-informaticiens devaient rejoindre l'agence.

Elle s'implique dans les actions d'animation scientifique du groupe DEBUG sur le site de Maisons-Alfort.

Elle contribue et participe à la réalisation de l'analyse quantitative des risques mise en place au sein de groupe de travail ou de saisine coordonnée par la DER. Son intervention pour les activités de référence (LNR, LRUE) ou les saisines s'effectuent sous la coordination du chef de la mission « coordination de la référence ».

3. Mode de fonctionnement

La mission travaille en mode projet. Une fiche est remplie par le porteur du projet. Une réunion permet d'évaluer les moyens techniques, humains à mettre en œuvre et de définir le niveau d'implication de la mission ainsi que l'horizon temporel des projets.

La mission est impliquée dans des projets portés par le laboratoire ou à l'Anses mais peut également développer des partenariats.

La mission est composée de quatre scientifiques, la publication scientifique représente une valorisation essentielle de ses activités. Le potentiel de valorisation des projets par les publications constitue un critère important à prendre en compte dans les phases initiales de discussion avec les porteurs de projet. On échangera également sur l'éventuelle présence des membres de la mission comme auteurs.

Une réunion des membres de la Mission est organisée bimensuellement.

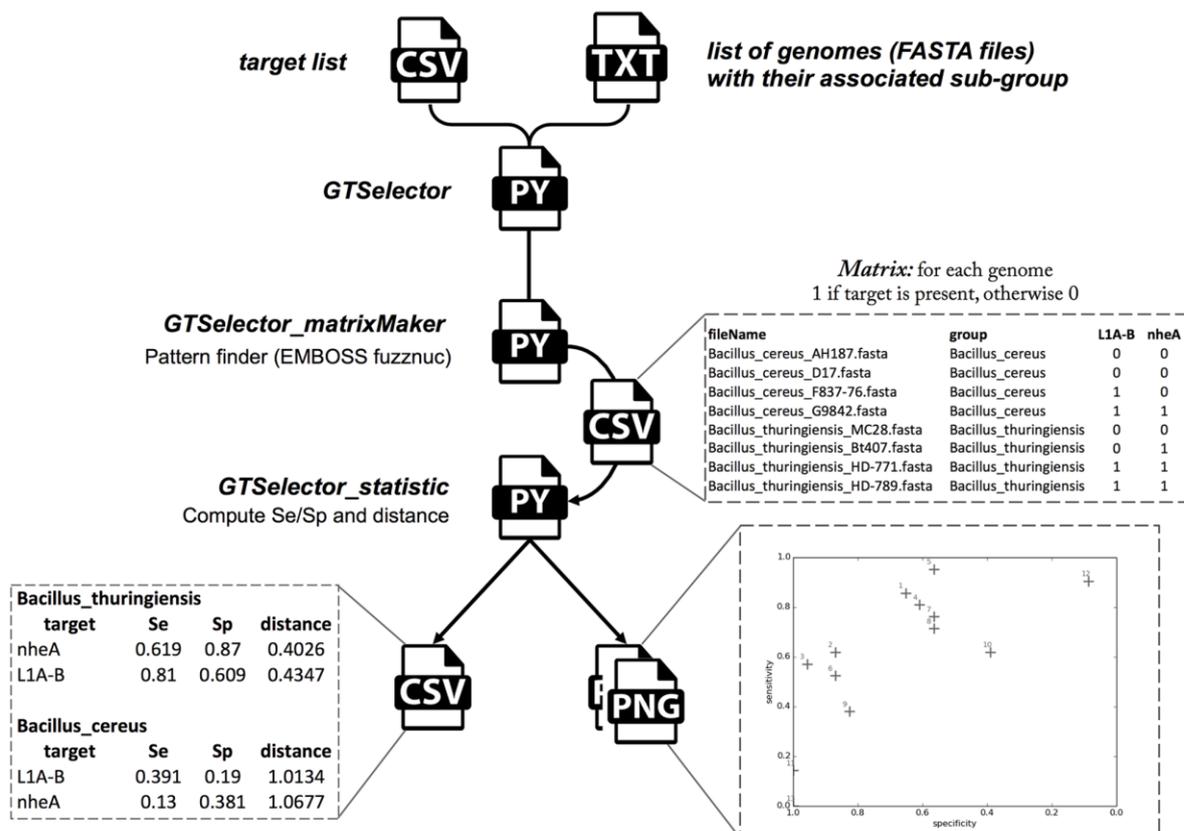
4. Organigramme fonctionnel

Cf. organigramme GAMeR

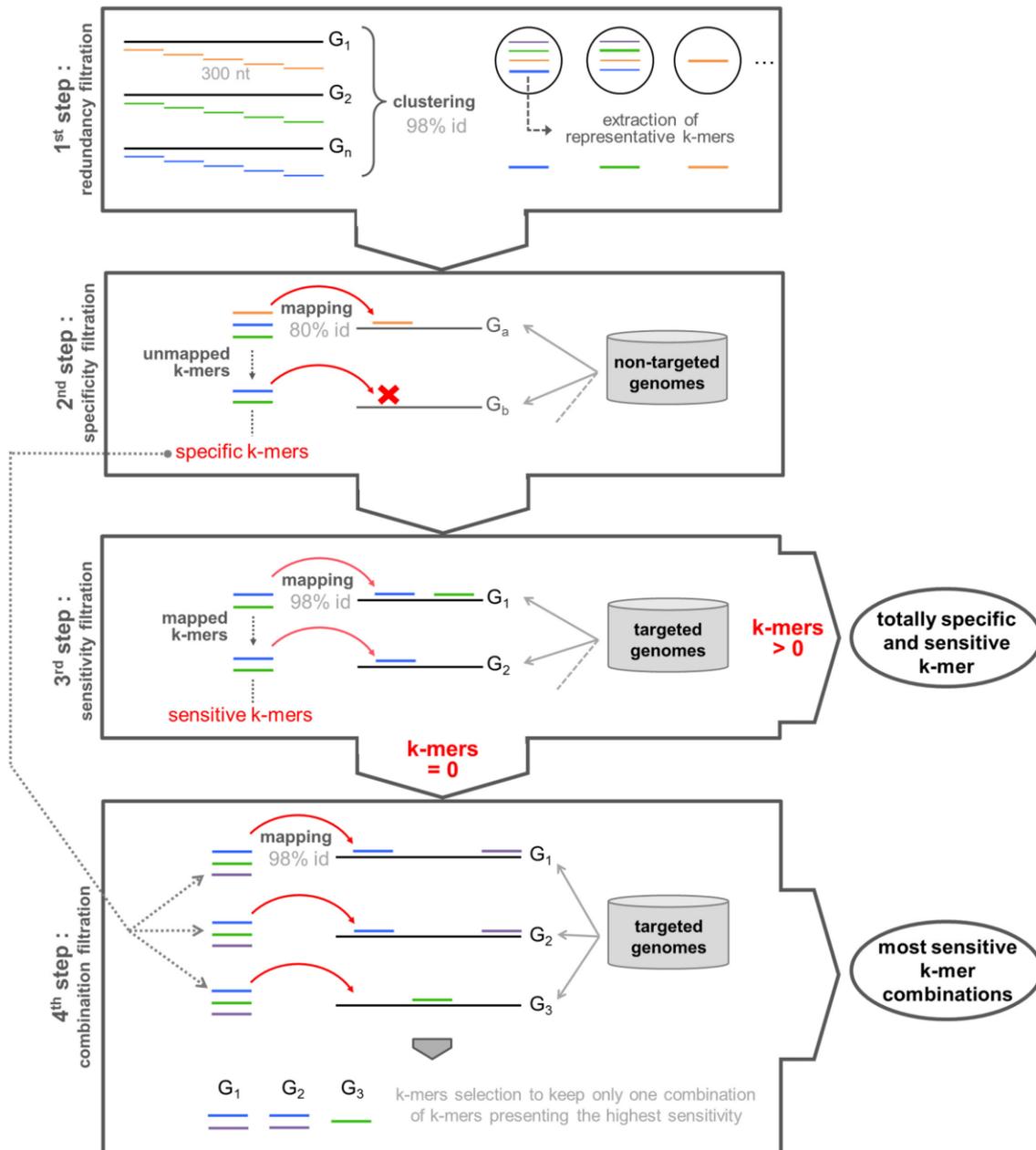
Anses LSAI - LSA n	NOTE D'ORGANISATION GAMeR LSA-NO-0023 (LSA-MM-0001 – Manuel Qualité A)	Révision	01	2/2
		Date	12/02/2016	

Seuls la version informatique et les exemplaires en diffusion contrôlée font foi.

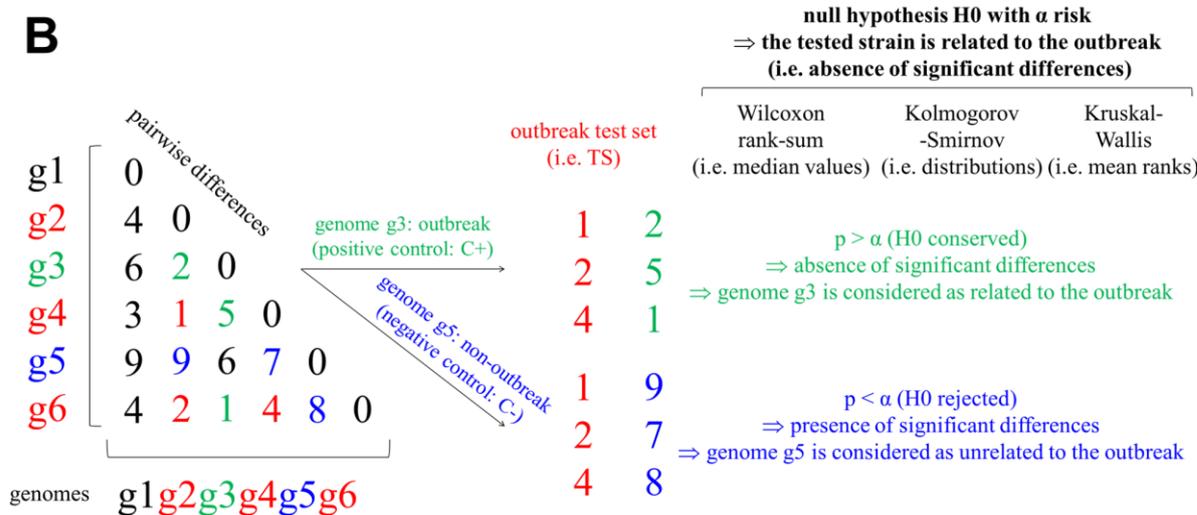
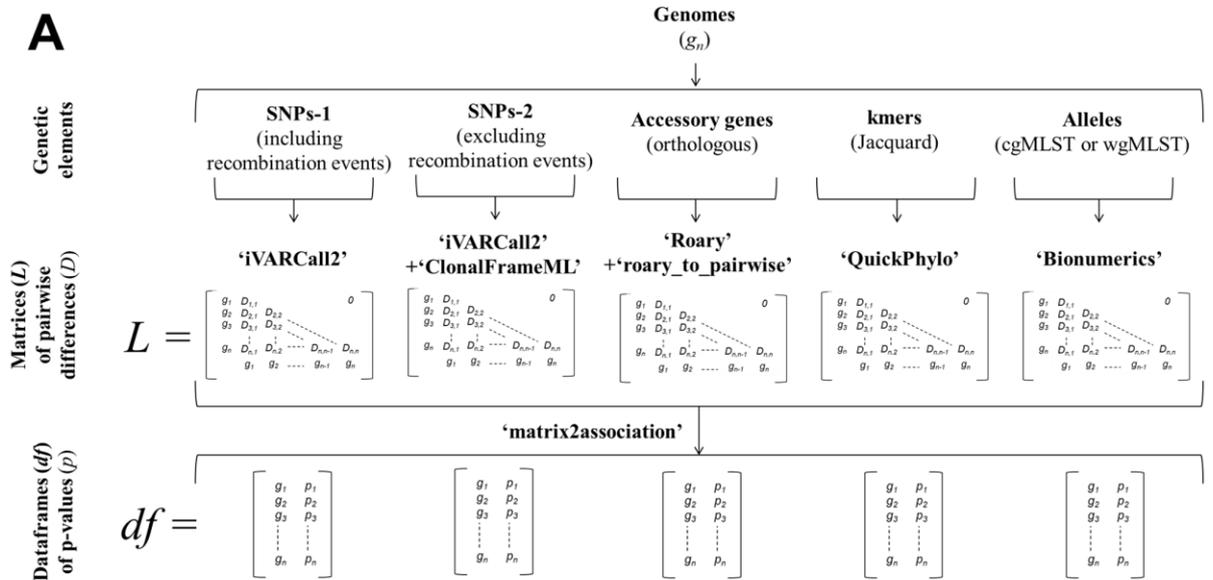
Annexe 7 : Outil « genotype evaluator » (GTEvaluator) développé par la mission génome analyse modélisation risque (GAMeR) au sein du laboratoire de sécurité des aliments (LSAL) pour estimer la justesse (spécificité et sensibilité) de cibles moléculaires à l'échelle génomique [95]



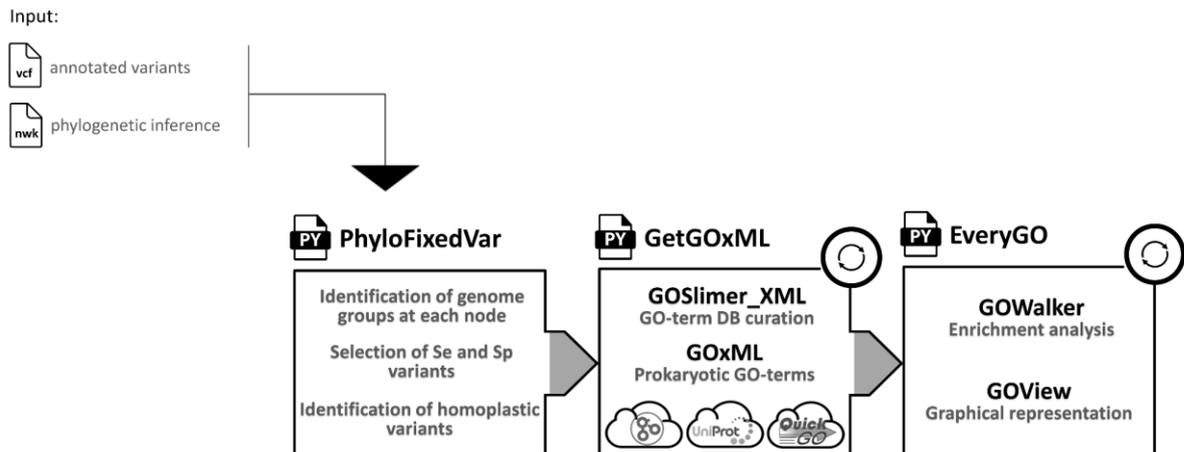
Annexe 8 : Outil « genotype finder » (GTFinder) développé par la mission génome analyse modélisation risque (GAMeR) au sein du laboratoire de sécurité des aliments (LSAL) pour identifier des kmers caractéristiques (spécifiques et sensibles) de sous-ensembles de génomes [91]



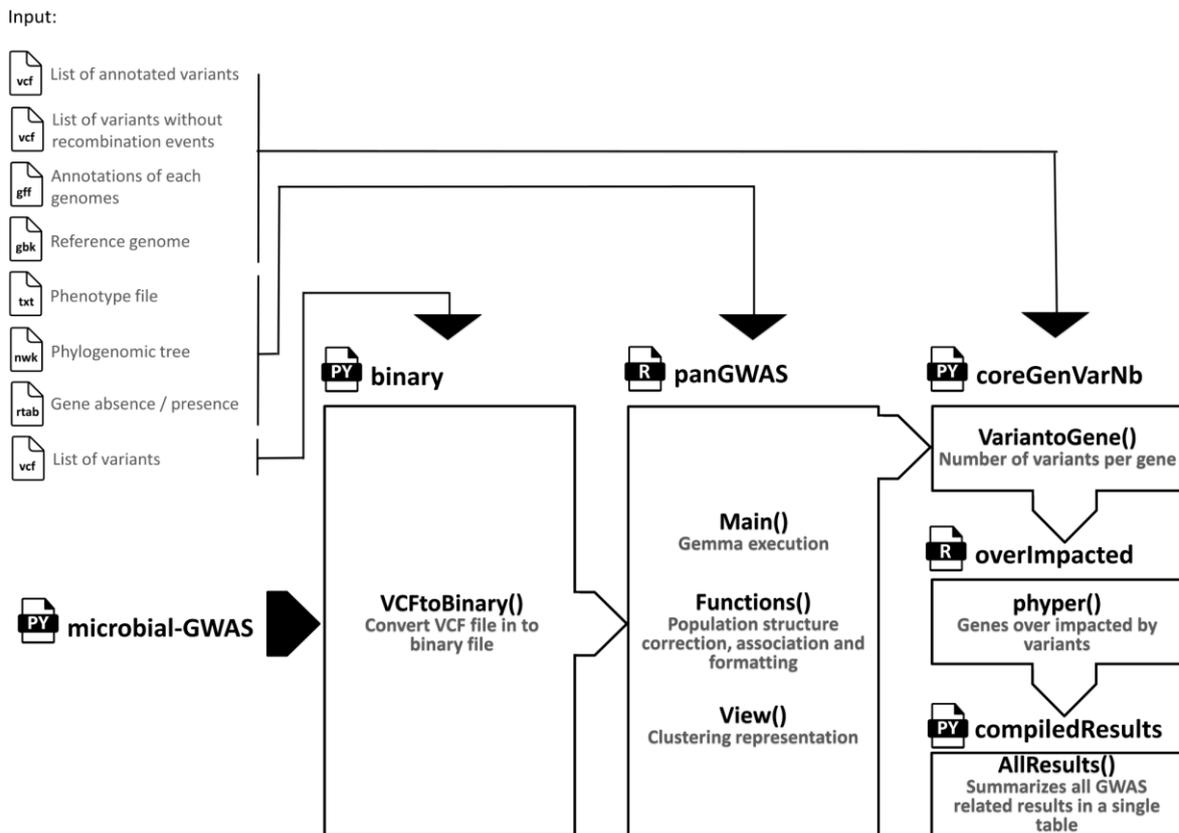
Annexe 9 : Outil (A) « Pairwise foodborne outbreaks » (Pairwise-FBO) et principe statistique (B) développés par la mission génome analyse modélisation risque (GAMeR) au sein du laboratoire de sécurité des aliments (LSAL) pour assigner des génomes à des toxi-infections alimentaires collectives (TIAC) sur la base de tests non-paramétriques (Kolmogorov-Smirnov, Wilcoxon-Mann-Whitney, Kruskal-Wallis) comparant les différences par paires de SNPs (avec ou sans recombinaisons : SNP-1 et SNP-2), gènes, kmers et allèles (cgMLST et wgMLST) [92]



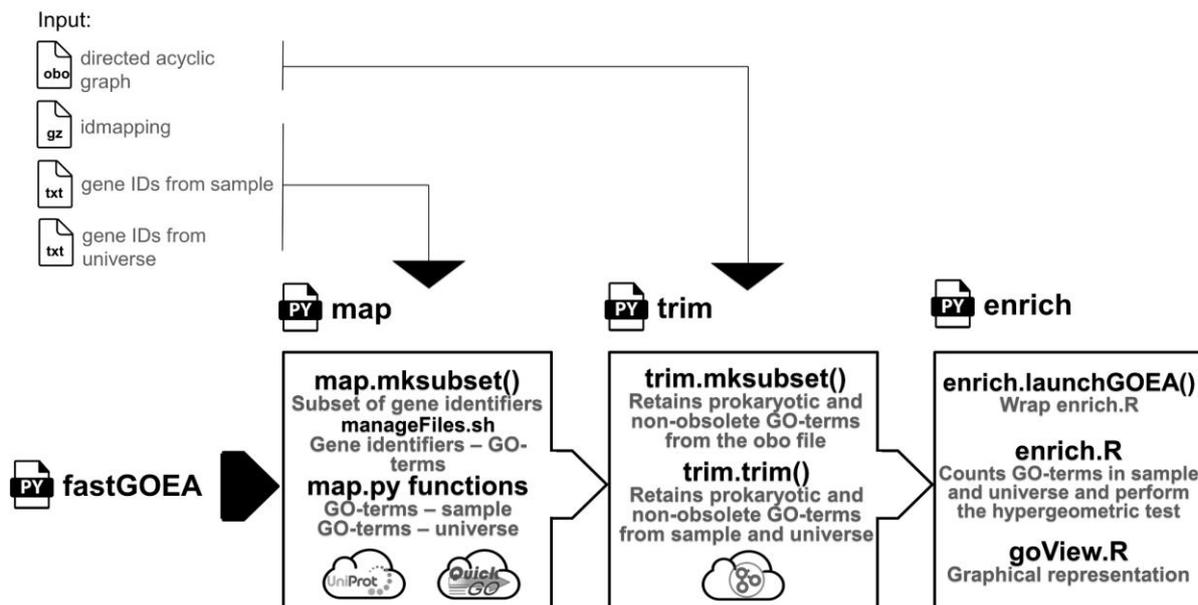
Annexe 10 : Outil « Phylogeny fixed variants » (PhyloFixedVar) développé par la mission génome analyse modélisation risque (GAMeR) au sein du laboratoire de sécurité des aliments (LSAL) afin d'identifier les variants du coregénom (SNPs and InDels), homoplastiques ou non-homoplastiques, fixés aux nœuds de l'inférence phylogénomique d'une population clonale ou panmictique, puis les voies métaboliques principalement enrichies par « gene ontology enrichment analysis » (GOEA) [101]



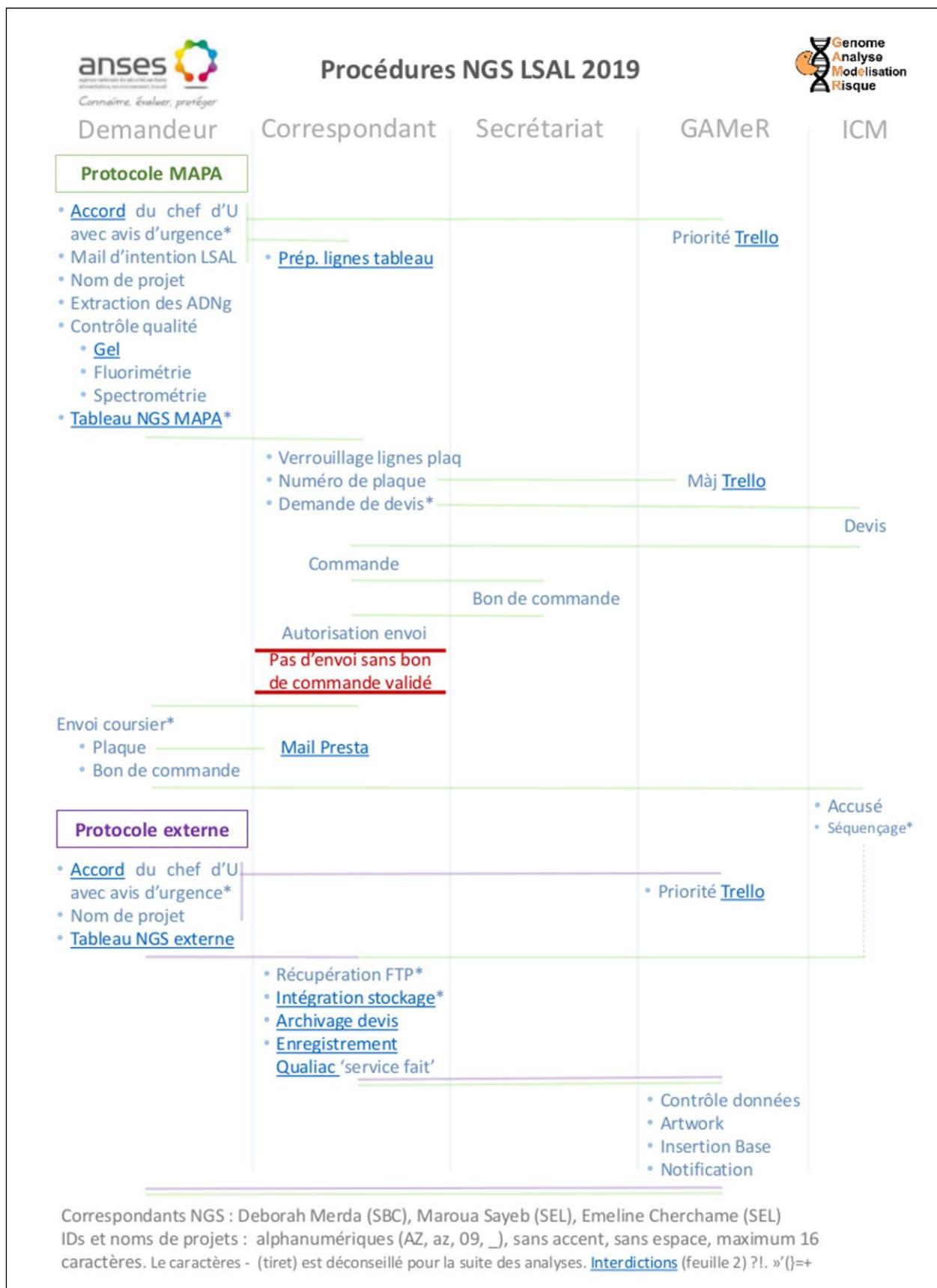
Annexe 11 : Outil « microbial genome wide association study » (microbial-GWAS) développé par la mission génome analyse modélisation risque (GAMeR) au sein du laboratoire de sécurité des aliments (LSAL) afin d'identifier les mutations (gènes accessoires et variants du coregénomme provenant d'événements de recombinaisons homologues ou non) associées à un trait phénotypique binaire dans une population clonale ou panmictique par « genome wide association study » (GWAS) [102]



Annexe 12 : Outil « fast gene ontology enrichment study » (fastGOES) développé par la mission génome analyse modélisation risque (GAMeR) au sein du laboratoire de sécurité des aliments (LSAL) afin d'identifier les voies métaboliques principalement enrichies par « gene ontology enrichment analysis » (GOEA) [102]



Annexe 13 : Procédures de soumission d'échantillons à la mission génome analyse modélisation risque (GAMeR) mises en place entre 2014 et 2019 pour appliquer le séquençage haut débit, l'assemblage *de novo* et l'identification de variants de génomes bactériens



Accord Chef d'Unité

Mail du chef d'U

Contenu: un bon pour accord, le nom du projet, le nombre d'échantillons, et le statut d'urgence si il y a lieu. Destinataires: membres GAMeR + demandeur + correspondant NGS.

Tableau NGS MAPA

Les données de **contrôles des extractions d'ADNg** sont enregistrées dans l'onglet '**Controls**' du fichier **NGS_MAPA.xlsx** qui est unique et incrémenté en échantillons.

Les images de gels sont stockées dans le dossier **NGS_gels** au format **.jpg** et nommées « **NuméroGel_InitialesOpérateur_DateExtractionAAA.MM.JJ_commentairePerso** » (<100 Ko par fichier). Le numéro de gel **DOIT** être sur **3 chiffres**, avec 0s si besoin. Ex: 003_CK_2019.02.21_test.jpg

Contrôle qualité et spécifications du partenaire de séquençage

Spécifications du partenaire de séquençage

- | | |
|--|---|
| <ul style="list-style-type: none"> - Contenant des échantillons : Plaques de 96 puits ou tubes - Eluant requis de reconstitution de l'ADN : Eau - Méthode de mesure de la concentration en ADN: Fluorimétrie - Quantité d'ADN minimale requise : 50ng - Concentration d'ADN minimale requise : Pas de prérequis | <ul style="list-style-type: none"> - Méthode de mesure des ratios de pureté de l'ADN : Spectrométrie - Pureté de l'ADN à 260/280 nm requise : Proche de 2.0 - Pureté de l'ADN à 260/230 nm requise : Proche de 2.0 - Preuve sur gel de l'intégrité de l'ADN extrait : Requis - Modalités d'expédition des échantillons : Echantillons réfrigérés et expédition par voie postale - Procédure de récupération des données : ftp |
|--|---|

Devis commande

- **un numéro de plaque** est assigné dans NGS_MAPA.xlsx
- **Demande de devis** émise au partenaire MAPA contenant gels + tableau compressés
- Une **commande** est formulée dès réception du devis

Envoi coursier

L'unité envoie par la poste (4°C) au partenaire de séquençage du MAPA à l'adresse suivante accompagnés du **bon de commande** incluant les **numéros de plaque** et de **devis**. L'envoi doit être accompagné d'un **mail à Yannick Marie** en demandant un **accusé de réception**

Pour Yannick Marie - Marché Anses MAPA-16A000026 - Plateforme de Génotypage-Séquençage - Institut du Cerveau et de la Moelle épinière (ICM) - Hôpital de la Salpêtrière - 47 boulevard de l'Hôpital - 75013 Paris.

AUCUNE PLAQUE NE DOIT ETRE ENVOYEE SANS BON DE COMMANDE

Séquençage 2-32 échantillons 5-10 jours

96-192 échantillons 10-15 jours

FTP et intégration

Intégration stockage : incorporation des **reads dans le système** soit dans le dossier « MAPA » soit « externe ». Dans MAPA, créer un sous-dossier portant le nom de plaque (**1 seul dossier par plaque**). Dans externe, créer un dossier de la date au format AAAA.MM.JJ

Un script réutilisant le tableau NGS (MAPA ou externe) et le dépôts de reads sera développé.

« Les gens font rarement ce en quoi ils croient. Ils font ce qui est convenable, puis ils regrettent »

Bob Dylan

RÉFÉRENCES

1. Cassaing J-C, Séramy P. Loi n°84-52 du 26 janvier 1984 sur l'enseignement supérieur.
2. Philip C, Girard J-F, Perget J. Arrêté du 23 novembre 1988 relatif à l'habilitation à diriger des recherches (NOR: MENU8802296A, Version consolidée au 12 juillet 2019).
3. Circulaire 89-004 du 5 janvier 1989. L'habilitation à diriger des recherches (RLR 430-5).
4. Circulaire du 27 octobre 1992. Circulaire relative au doctorat et à l'habilitation à diriger des recherches (Ministère de l'éducation nationale et de la culture 27 oct 1992).
5. Potier P. Arrêté du 13 juillet 1995 modifiant l'arrêté du 23 novembre 1988 relatif à l'habilitation à diriger des recherches (JORF n°171 du 25 juillet 1995 page 11019, NOR: MENH9501401A).
6. Lang J, Schwartzberg R-G, Kouchner B. Arrêté du 25 avril 2002 modifiant l'arrêté du 23 novembre 1988 relatif à l'habilitation à diriger des recherches (JORF n°99 du 27 avril 2002 page 7635 texte n° 59, NOR: MENS0200985A).
7. Vallaud-Belkacem N, Mandon T. Arrêté du 25 mai 2016 fixant le cadre national de la formation et les modalités conduisant à la délivrance du diplôme national de doctorat (JORF n°0122 du 27 mai 2016, texte n° 10, NOR: MENS1611139A).
8. Bonnafous S. Arrêté du 1er juillet 2016 modifiant l'arrêté du 25 mai 2016 fixant le cadre national de la formation et les modalités conduisant à la délivrance du diplôme national de doctorat (JORF n°0173 du 27 juillet 2016 texte n° 8, NOR: MENS1615695A).
9. Courtillot V, Decomps B. Arrêté du 13 février 1992 modifiant l'arrêté du 23 novembre 1988 relatif à l'habilitation à diriger des recherches (JORF n°44 du 21 février 1992 page 2697, NOR: MENH9102942A).
10. de Robien G, Goulard F. Arrêté du 7 août 2006 relatif à la formation doctorale (NOR: MENS0602083A).
11. Procédure Université Paris-Est. Habilitation à diriger des recherches - Procédure d'inscription et de soutenance. Validé 15 Juin 2018. :1–4.
12. Boutibonnes P. L'œil de Leeuwenhoek et l'invention de la microscopie. *Alliage*. 1999;39 ISSN 1144-5645:58–66.
13. Finlay BJ, Esteban GF. Exploring Leeuwenhoek's legacy: the abundance and diversity of protozoa. *Int Microbiol*. 2001;4 ISSN 1139-6709:125–31.
14. Hamraoui É. Van Leeuwenhoek Antonie, 1632-1723. *Dict Hist Philos Sci LECOURT Dir*. 1999;Presses universitaires de France Paris:1–970.
15. Rostand J. La Genèse de la vie. Histoire des idées sur la génération spontanée. Hachette. 1943;Paris:1–205.
16. Van Tieghem PÉ. Identité du *Bacillus amylobacter* et du vibrion butyrique de M. Pasteur. *Comtes Rendus Académie Sci*. 1879;t:5–8.

17. Pasteur L. Mémoire sur la relation qui peut exister entre la forme cristalline et la composition chimique, et sur la cause de la polarisation rotatoire (Extrait par l'auteur). Comptes Rendus Académie Sci Séance 22 Mai 1848. 1848;t:535–8.
18. Pasteur L. Études sur le vin, ses maladies, causes qui les provoquent, procédés nouveaux pour le conserver et pour le vieillir. Imprimerie impériale. Paris; 1866.
19. Pasteur L. Études sur la maladie des vers à soie. Œuvres Complètes Pasteur. 1870;t:166–7.
20. Debré P. Louis Pasteur. Gd Livre Mois. 1994;in-8:1–563.
21. Chamberland C. Le charbon et la vaccination charbonneuse d'après les travaux récents de M. Pasteur. Bernard Tignol Paris. 1883;17 janvier:1–314.
22. Tigertt WD. Anthrax. William Smith Greenfield, M.D., F.R.C.P., Professor Superintendent, the Brown Animal Sanatory Institution (1878-1881). Concerning the priority due to him for the production of the first vaccine against anthrax. J Hyg (Lond). 1980;85:415–20.
23. Bazin H. Histoire des vaccinations. John Libbey Eurotext Paris. 2008;ISBN-10: 2742007059 ISBN-13: 978-2742007059:1–471.
24. Pasteur L. Méthode pour prévenir la rage après morsure. Comptes Rendus Académie Sci Séance 26 Octobre 1885. 1885;t:765–74.
25. Ullmann A. Pasteur–Koch: Distinctive ways of thinking about infectious diseases. Microbe. 2007;2:383–7.
26. Koch R. Die Ätiologie der Tuberkulose. Berl Klin Wochen- Schr. 1882;10. April Nr. 15 cf:428–45.
27. Howard-Jones N. Robert Koch and the cholera vibrio: A centenary. Br Med J (Clin Res Ed). 1984;288:379–81.
28. Fanica P-O. Le lait la vache et le citadin. Du XVIIe au XXe siècle. Éditions Quae. 2008;1:1–490.
29. Lagrange E. Robert Koch, sa vie et son œuvre. Rev Hist Sci. 1948;1:278–9.
30. Campbell WC. Jacob Henle. Well-known teacher of Koch and progenitor of his Postulates. Drew Univ Madison. NJ 07940 Germ Theory Calendar.
31. Koch R. Weitere Mitteilungen uber ein Heilmittel gegen Tuberkulose. Robert Koch-Inst. 2010;:661–8.
32. Manwaring WH. Fleming's "Lysozyme." Calif West Med. 1942;56:5.
33. Roussel C, Vial F, Heymans G, Rullière R. Des moisissures à la pénicilline Quelques "prélèvements" dans la "colonie" des précurseurs. Société Fr Médecine Séance 24 Janvier 1981. 1981;Chaire d'histoire de la médecine Paris-VI:29–38.
34. Luria SE, Delbrück M. Mutations of Bacteria from Virus Sensitivity to Virus Resistance. Genetics. 1943;28:491–511.

35. Watson JD, Crick FHC. A structure for deoxyribose nucleic acid *Nature*. *Nature*. 1953;171:737–8.
36. Mullis KB. The unusual origin of the polymerase chain reaction. *Sci Am*. 1990;262:56–61, 64–5.
37. Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, et al. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*. 1988;239:487–91.
38. Comas I, Coscolla M, Luo T, Borrell S, Holt KE, Kato-Maeda M, et al. Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. *Nat Genet*. 2013;45:1176–82.
39. Stadtländer CTK-H. One Health: people, animals, and the environment. *Infect Ecol Epidemiol*. 2015;5:30514.
40. Alibek K. Smallpox: a disease and a weapon. *Int J Infect Dis*. 2004;8:3–8.
41. Furuse Y, Suzuki A, Oshitani H. Origin of measles virus: divergence from rinderpest virus between the 11th and 12th centuries. *Virology*. 2010;7:52.
42. de Romilly J, Weil R, Bodin L. Thucydide, Histoire de la guerre du Péloponnèse. Robert Laffont Paris. 1990;ISBN-10: 2221058542 ISBN-13: 978-2221058541:1–840.
43. Monceaux P, Sebaï L. Les Intellectuels carthaginois (Les Africains). Éditions Cartaginoiseries Tunis. 2009;paysages littéraires méditerranéens ISBN : 978-9973-704-09-2:1–168.
44. Harbeck M, Seifert L, Hänsch S, Wagner DM, Birdsell D, Parise KL, et al. *Yersinia pestis* DNA from skeletal remains from the 6th century AD reveals insights into Justinianic plague. *PLoS Pathog*. 2013;9:e1003349.
45. de Lannoy F. Pestes et épidémies au Moyen Age (VIe-XVe siècles). Ouest Fr. ISBN 978-2-7373-6719-9 Grand Format:1–128.
46. Luciani D. Le Lévitique. Éthique & esthétique (coll. Connaître la Bible). Rev Théologique Louvain Année 2006. 2006;37:418–9.
47. Krause-Kyora B, Nutsua M, Boehme L, Pierini F, Pedersen DD, Kornell S-C, et al. Ancient DNA study reveals HLA susceptibility locus for leprosy in medieval Europeans. *Nat Commun*. 2018;9:1569.
48. Schuenemann VJ, Kumar Lankapalli A, Barquera R, Nelson EA, Iraíz Hernández D, Acuña Alonzo V, et al. Historic *Treponema pallidum* genomes from colonial Mexico retrieved from archaeological remains. *PLoS Negl Trop Dis*. 2018;12:e0006447.
49. Cadotte M, James-abra E. Épidémie. *Encycl Can*. 2015;en ligne le 19 juin 2013 modifié le 23 juillet 2015:1.
50. Azizi M, Azizi F. History of cholera outbreaks in Iran during the 19(th) and 20(th) Centuries. *Middle East J Dig Dis*. 2010;2:51–5.

51. Nolan CM, Elarth AM, Barr H, Saeed AM, Risser DR. An outbreak of tuberculosis in a shelter for homeless men: A description of its evolution and control. *Am Rev Respir Dis.* 1991;143:257–61.
52. Trentini F, Poletti P, Baldacchino F, Drago A, Montarsi F, Capelli G, et al. The containment of potential outbreaks triggered by imported Chikungunya cases in Italy: a cost utility epidemiological assessment of vector control measures. *Sci Rep.* 2018;8:9034.
53. Olivero J, Fa JE, Real R, Márquez AL, Farfán MA, Vargas JM, et al. Recent loss of closed forests is associated with Ebola virus disease outbreaks. *Sci Rep.* 2017;7:14291.
54. Hajra A, Bandyopadhyay D, Hajra S. Zika virus: A global threat to humanity: A comprehensive review and current developments. *North Am J Med Sci.* 2016;8:123.
55. Chowell G, Castillo-Chavez C, Fenimore PW, Kribs-Zaleta CM, Arriola L, Hyman JM. Model parameters and outbreak control for SARS. *Emerg Infect Dis.* 2004;10:1258–63.
56. Chatziprodromidou IP, Arvanitidou M, Guitian J, Apostolou T, Vantarakis G, Vantarakis A. Global avian influenza outbreaks 2010–2016: a systematic review of their distribution, avian species and virus subtype. *Syst Rev.* 2018;7:17.
57. Anyamba A, Chretien J-P, Britch SC, Soebiyanto RP, Small JL, Jepsen R, et al. Global disease outbreaks associated with the 2015–2016 El Niño event. *Sci Rep.* 2019;9:1930.
58. Golay C. The food crisis and food security: Towards a new world food order? *Rev Int Polit Dév.* 2010;1:215–32.
59. Zhao D, Liu R, Zhang X, Li F, Wang J, Zhang J, et al. Replication and virulence in pigs of the first African swine fever virus isolated in China. *Emerg Microbes Infect.* 2019;8:438–47.
60. Hodges JR, Kimball A. Globalization and health. *Glob Health.* 2005;1:1–4.
61. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S Am.* 1977;74:5463–7.
62. Maxam AM, Gilbert W. A new method for sequencing DNA. *Proc Natl Acad Sci.* 1977;74:560–4.
63. Jorgenson JW, Lukacs KD. Zone electrophoresis in open-tubular glass capillaries. *Anal Chem.* 1981;53:1298–302.
64. Karger BL, Guttman A. DNA sequencing by CE. *Electrophoresis.* 2009;30:S196–202.
65. O'Connor M, Peifer M, Bender W. Construction of large DNA segments in *Escherichia coli*. *Science.* 1989;244:1307–12.
66. Zhang J-H, Wu L-Y, Zhang X-S. Reconstruction of DNA sequencing by hybridization. *Bioinformatics.* 2003;19:14–21.
67. Istrail S, Sutton GG, Florea L, Halpern AL, Mobarry CM, Lippert R, et al. Whole-genome shotgun assembly and comparison of human genome assemblies. *Proc Natl Acad Sci.* 2004;101:1916–21.

68. Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, et al. The complete genome of an individual by massively parallel DNA sequencing. *Nature*. 2008;452:872–6.
69. Levy S, Sutton G, Ng PC, Feuk L, Halpern AL, Walenz BP, et al. The diploid genome sequence of an individual Human. *PLoS Biol*. 2007;5:e254.
70. Besser J, Carleton HA, Gerner-Smidt P, Lindsey RL, Trees E. Next-generation sequencing technologies and their application to the study and control of bacterial infections. *Clin Microbiol Infect*. 2018;24:335–41.
71. Stark Z, Dolman L, Manolio TA, Ozenberger B, Hill SL, Caulfield MJ, et al. Integrating genomics into healthcare: A global responsibility. *Am J Hum Genet*. 2019;104:13–20.
72. Karsch-Mizrachi I, Takagi T, Cochrane G, on behalf of the International Nucleotide Sequence Database Collaboration. The international nucleotide sequence database collaboration. *Nucleic Acids Res*. 2018;46:D48–51.
73. ten Hoopen P, Amid C, Luigi Buttigieg P, Pafilis E, Bravakos P, Cerdeño-Tárraga AM, et al. Value, but high costs in post-deposition data curation. *Database*. 2016;2016:bav126.
74. Beagrie N, Curatcha J. The value and impact of the european bioinformatics institute. project report. Charles Beagrie Ltd Hinxton UK. 2016;:1–96.
75. Wang Y, Song F, Zhu J, Zhang S, Yang Y, Chen T, et al. GSA: Genome Sequence Archive *. *Genomics Proteomics Bioinformatics*. 2017;15:14–8.
76. Ellenberg J, Swedlow JR, Barlow M, Cook CE, Sarkans U, Patwardhan A, et al. A call for public archives for biological image data. *Nat Methods*. 2018;15:849–54.
77. Birney E, Vamathevan J, Goodhand P. Genomics in healthcare: GA4GH looks to 2022. 2017. doi:10.1101/203554.
78. Karsenti E, Acinas SG, Bork P, Bowler C, De Vargas C, Raes J, et al. A holistic approach to marine eco-systems biology. *PLoS Biol*. 2011;9:e1001177.
79. Allard MW, Strain E, Melka D, Bunning K, Musser SM, Brown EW, et al. Practical value of food pathogen traceability through building a whole-genome sequencing network and database. *J Clin Microbiol*. 2016;54:1975–83.
80. ECDC-EFSA, Van Walle I, Guerra B, Borges V, André Carriço J, Guy C, et al. EFSA and ECDC technical report on the collection and analysis of whole genome sequencing data from food-borne pathogens and other relevant microorganisms isolated from human, animal, food, feed and food/feed environmental samples in the joint ECDC-EFSA molecular typing database. *EFSA Support Publ*. 2019;EN:1–92.
81. Grant K, Jenkins C, Arnold C, Green J, Zambon M. Implementing pathogen genomics: A case study. *Public Health Engl - Prot Improv Nations Health*. phg foundation making science work for health PHE publications gateway number: 2018254:1–32.
82. Dragacci S. Passer son HDR (décembre 2016) - Mon HDR en mode projet. *Anses Procédure*. 1:1–6.

83. Sarkozy N, Fillon F, Borloo J-L, Lagarde C, Woerth E, Bachelot-Narquin R, et al. Ordonnance n° 2010-18 du 7 janvier 2010 portant création d'une agence nationale chargée de la sécurité sanitaire de l'alimentation, de l'environnement et du travail. 2010.
84. CDDEP. Charte développement durable des établissements publics et entreprises publiques. Ministère Écologie Dév Durable Transp Logement. 2008;1:1–2.
85. Genet R, Toulhoat P, Jacquot-Guimbal H, Cointe R, Niel J-C, Bournigal J-M, et al. Charte de l'ouverture à la société des organismes publics de recherche, d'expertise et d'évaluation des risques sanitaires et environnementaux. Anses. 2016;1:1–4.
86. Travert S, Mauguin P, Genet R. Prévention des risques sanitaires : Stéphane Travert, ministre de l'Agriculture et de l'Alimentation renforce l'épidémiosurveillance en partenariat avec l'INRA et l'Anses. Commun Press. 2018;1:1–2.
87. EFSA-ECDC. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2017. Eur Food Saf Auth J. 2018;16:5500.
88. EFSA-ECDC. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2015. Eur Food Saf Auth J. 2016;14.
89. EFSA-ECDC. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2016. Eur Food Saf Auth J. 2017;15:5077.
90. WHO. Whole genome sequencing for foodborne disease surveillance. World Health Organ Landsc Pap. 2018;ISBN:1–54.
91. Cadel-Six S, Felten A, Fach P, Guillier L, Radomski N, Mistou M-Y, et al. Algorithme de sélection de kmers sensibles et spécifique à l'échelle génomique et vérification des cibles moléculaire par PCR haut débit pour identifier les sérovars de *Salmonella*. 2018.
92. Radomski N, Cadel-Six S, Cherchame E, Barbet P, Palma F, Mallet L, et al. A simple and robust statistical method to define genetic relatedness of samples related to outbreaks at the genomic scale - Application to retrospective *Salmonella* foodborne outbreak investigations. Front Microbiol. 2019;10:1–13.
93. Mughini-Gras L, Kooh P, Augustin J-C, David J, Fravalo P, Guillier L, et al. Source attribution of foodborne diseases: potentialities, hurdles, and future expectations. Front Microbiol. 2018;9. doi:10.3389/fmicb.2018.01983.
94. Henri C, Leekitcharoenphon P, Carleton HA, Radomski N, Kaas RS, Mariet J-F, et al. An assessment of different genomic approaches for inferring phylogeny of *Listeria monocytogenes*. Front Microbiol. 2017;8. doi:10.3389/fmicb.2017.02351.
95. Felten A, Guillier L, Radomski N, Mistou M-Y, Lailler R, Cadel-Six S. Genome Target Evaluator (GTEvaluator): A workflow exploiting genome dataset to measure the sensitivity and specificity of genetic markers. PLOS ONE. 2017;12:e0182082.
96. Sévellec Y, Vignaud M-L, Granier SA, Lailler R, Feurer C, Le Hello S, et al. Polyphyletic nature of *Salmonella enterica* serotype Derby and lineage-specific host-association revealed by genome-wide analysis. Front Microbiol. 2018;9. doi:10.3389/fmicb.2018.00891.

97. Mahamat Abdelrahim A, Radomski N, Delannoy S, Djellal S, Le Négrate M, Hadjab K, et al. Large-scale genomic analyses and toxinotyping of *Clostridium perfringens* implicated in foodborne outbreaks in France. *Front Microbiol.* 2019;10:777.
98. Sévellec Y, Felten A, Radomski N, Granier S, Le Hello S, Petrovska L, et al. Genetic diversity of *Salmonella* Derby from the poultry sector in Europe. *Pathogens.* 2019;8:46.
99. Fritsch L, Felten A, Palma F, Mariet J-F, Radomski N, Mistou M-Y, et al. Insights from genome-wide approaches to identify variants associated to phenotypes at pan-genome scale: Application to *L. monocytogenes*' ability to grow in cold conditions. *Int J Food Microbiol.* 2019;291:181–8.
100. Sévellec Y, Granier SA, Radomski N, Felten A, Le Hello S, Feurer C, et al. Complete genome sequence of *Salmonella enterica* subsp. *enterica* serotype Derby, Associated with the pork sector in France. *Microbiol Resour Announc.* 2018;7. doi:10.1128/MRA.01027-18.
101. Felten A, Vila Nova M, Durimel K, Guillier L, Mistou M-Y, Radomski N. First gene-ontology enrichment analysis based on bacterial coregenome variants: insights into adaptations of *Salmonella* serovars to mammalian- and avian-hosts. *BMC Microbiol.* 2017;17:1–22.
102. Vila Nova M, Durimel K, La K, Felten A, Bessières P, Mistou M-Y, et al. Genetic and metabolic signatures of *Salmonella enterica* subsp. *enterica* associated with animal sources at the pangenomic scale. *BMC Genomics.* 2019;20:1–21.
103. Douarre P-E, Mallet L, Radomski N, Felten A, Mistou M-Y. Analysis of COMPASS, a new comprehensive plasmid database. *Submiss.* 2020.
104. Bonis M, Felten A, Payraud S, Dijoux A, Mallet L, Radomski N, et al. Characterization of *Bacillus thuringiensis* isolates associated to foodborne outbreaks in France from 2007 to 2017. *Submiss.* 2020.
105. Jolley KA, Maiden MC. BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics.* 2010;11:595.
106. Moura A, Criscuolo A, Pouseele H, Maury MM, Leclercq A, Tarr C, et al. Whole genome-based population biology and epidemiological surveillance of *Listeria monocytogenes*. *Nat Microbiol.* 2016;2:16185.
107. Alikhan N-F, Zhou Z, Sergeant MJ, Achtman M. A genomic overview of the population structure of *Salmonella*. *PLOS Genet.* 2018;14:e1007261.
108. Timme RE, Sanchez Leon M, Allard MW. Utilizing the public GenomeTrakr database for foodborne pathogen traceback. In: Bridier A, editor. *Foodborne Bacterial Pathogens.* New York, NY: Springer New York; 2019. p. 201–12. doi:10.1007/978-1-4939-9000-9_17.
109. Llarena A-K, Ribeiro-Gonçalves BF, Nuno Silva D, Halkilahti J, Machado MP, Santos Da Silva M, et al. INNUENDO: A cross-sectoral platform for the integration of genomics in the surveillance of food-borne pathogens. *EFSA Support Publ.* 2018;EN:1–142.
110. Matthews TC, Bristow FR, Griffiths EJ, Petkau A, Adam J, Dooley D, et al. The integrated rapid infectious disease analysis (IRIDA) platform. preprint. *Bioinformatics;* 2018. doi:10.1101/381830.

111. Argimón S, Abudahab K, Goater RJE, Fedosejev A, Bhai J, Glasner C, et al. Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. *Microb Genomics*. 2016;2. doi:10.1099/mgen.0.000093.
112. Jolley KA, Bray JE, Maiden MCJ. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res*. 2018;3:124.
113. Nadon C, Van Walle I, Gerner-Smidt P, Campos J, Chinen I, Concepcion-Acevedo J, et al. PulseNet International: Vision for the implementation of whole genome sequencing (WGS) for global food-borne disease surveillance. *Eurosurveillance*. 2017;22. doi:10.2807/1560-7917.ES.2017.22.23.30544.
114. Schmid D, Allerberger F, Huhulescu S, Pietzka A, Amar C, Kleta S, et al. Whole genome sequencing as a tool to investigate a cluster of seven cases of listeriosis in Austria and Germany, 2011-2013. *Clin Microbiol Infect Off Publ Eur Soc Clin Microbiol Infect Dis*. 2014;20:431–6.
115. Radomski N. Sources des mycobactéries non-tuberculeuses dans les bassins versants. Thèse Univ Paris-Est Éc Ponts ParisTech. 2011;École Doctorale : Sciences, Ingénierie et Environnement Directeur de thèse : Moilleron, R. Co-encadrante de thèse : Lucas, F.S.:1–220.
116. Radomski N, Moilleron R, Lucas FS, Falkinham III JO. Challenges in environmental monitoring of pathogens: Case study in *Mycobacterium avium*. In: Méndez-Vilas A, editor. *Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology*. Badajoz: Formatex Research Center; 2010. p. 1551–61.
117. Midelet-Bourdin G, Copin S, Leleu G, Malle P. Determination of *Listeria monocytogenes* growth potential on new fresh salmon preparations. *Food Control*. 2010;21:1415–8.
118. Lee RS, Radomski N, Proulx J-F, Levade I, Shapiro BJ, McIntosh F, et al. Population genomics of *Mycobacterium tuberculosis* in the Inuit. *Proc Natl Acad Sci U S A*. 2015;112:13609–14.
119. Wang J, McIntosh F, Radomski N, Dewar K, Simeone R, Enninga J, et al. Insights on the emergence of *Mycobacterium tuberculosis* from the analysis of *Mycobacterium kansasii*. *Genome Biol Evol*. 2015;7:856–70.
120. Lee RS, Radomski N, Proulx JF, Manry J, McIntosh F, Desjardins F, et al. Re-emergence and amplification of tuberculosis in the Canadian Arctic. *J Infect Dis*. 2015;pii:1–10.
121. Domenech P, Rog A, Moolji J, Radomski N, Fallow A, Leon-Solis L, et al. The origins of a 350-kilobase genomic duplication in *Mycobacterium tuberculosis* and its impact on virulence. *Infect Immun*. 2014;58:6962–5.
122. Radomski N, Roguet A, Lucas FS, Veyrier FJ, Cambau E, Accrombessi H, et al. *atpE* gene as a new useful specific molecular target to quantify *Mycobacterium* in environmental samples. *BMC Microbiol*. 2013;13:277.
123. Radomski N, Kreitmann L, McIntosh F, Behr MA. The critical role of DNA extraction for detection of mycobacteria in tissues. *PLoS One*. 2013;8:e78749.

124. Radomski N, Betelli L, Moilleron R, Haenn S, Moulin L, Cambau E, et al. *Mycobacterium* behavior in wastewater treatment plant, a bacterial model distinct from *Escherichia coli* and enterococci. *Environ Sci Technol*. 2011;45:5380–6.
125. Radomski N, Lucas FS, Moilleron R, Cambau E, Haenn S, Moulin L. Development of a real-time qPCR method for detection and enumeration of *Mycobacterium* spp. in surface water. *Appl Environ Microbiol*. 2010;76:7348–51.
126. Radomski N, Cambau E, Moulin L, Haenn S, Moilleron R, Lucas FS. Comparison of culture methods for isolation of nontuberculous mycobacteria from surface waters. *Appl Environ Microbiol*. 2010;76:3514–3520.
127. Radomski N, Thibault VC, Karoui C, de Cruz K, Cochard T, Gutiérrez C, et al. Genotypic diversity of *Mycobacterium avium* subspecies from human and animal origins, studied by MIRU-VNTR and *IS1311* RFLP typing methods. *J Clin Microbiol*. 2010;48:1026–1034.
128. Palma F, Brauge T, Radomski N, Mallet L, Felten A, Mistou M-Y, et al. Dynamics of mobile genetic elements of *Listeria monocytogenes* persisting in ready-to-eat seafood processing plants in France. *BMC Genomics*. 2020;21:130.
129. DGMDP Direction générale de la mondialisation, du développement et des partenariats. Position française sur le concept « One Health/Une seule santé »; 2011:1-32.
130. projet de la loi Elan. LOI n° 2018-1021 du 23 novembre 2018 portant évolution du logement, de l'aménagement et du numérique - JORF n°0272 du 24 novembre 2018 - NOR: TERL1805474L.
131. Food Safety WGS2019. Foodborne pathogens and whole genome sequencing: impact on public health protection. Maison de la RATP - Espace du Centenaire 189 rue de Bercy 75012 Paris - France; 2019. p. 1–60.
132. Frank D. One world, one health, one medicine. *Can Vet J Rev Veterinaire Can*. 2008;49:1063–5.
133. Osburn B, Scott C, Gibbs P. One world--one medicine--one health: emerging veterinary challenges and opportunities. *Rev Sci Tech Int Off Epizoot*. 2009;28:481–6.
134. van Helden PD, van Helden LS, Hoal EG. One world, one health. Humans, animals and the environment are inextricably linked--a fact that needs to be remembered and exploited in our modern approach to health. *EMBO Rep*. 2013;14:497–501.
135. Angelos J, Arens A, Johnson H, Cadriel J, Osburn B. One Health in food safety and security education: A curricular framework. *Comp Immunol Microbiol Infect Dis*. 2016;44:29–33.
136. Le VTM, Diep BA. Selected insights from application of whole-genome sequencing for outbreak investigations: *Curr Opin Crit Care*. 2013;19:432–9.
137. Gantois I, Ducatelle R, Pasmans F, Haesebrouck F, Gast R, Humphrey TJ, et al. Mechanisms of egg contamination by *Salmonella* Enteritidis. *FEMS Microbiol Rev*. 2009;33:718–38.

138. Gast RK, Guraya R, Jones DR, Anderson KE. Contamination of eggs by *Salmonella* Enteritidis in experimentally infected laying hens housed in conventional or enriched cages. *Poult Sci.* 2014;93:728–33.
139. EFSA. Assessing determinants of the non-decreasing incidence of *Salmonella* Enteritidis. *EFSA J.* 2019.
140. Zhang S, Li S, Gu W, den Bakker H, Boxrud D, Taylor A, et al. Zoonotic source attribution of *Salmonella enterica* serotype Typhimurium using genomic surveillance data, United States. *Emerg Infect Dis.* 2019;25:82–91.
141. Liu H, Whitehouse CA, Li B. Presence and persistence of *Salmonella* in water: The impact on microbial quality of water and food safety. *Front Public Health.* 2018;6:159.
142. Melotto M, Panchal S, Roy D. Plant innate immunity against human bacterial pathogens. *Front Microbiol.* 2014;5:1–12.
143. Greene SK, Daly ER, Talbot EA, Demma LJ, Holzbauer S, Patel NJ, et al. Recurrent multistate outbreak of *Salmonella* Newport associated with tomatoes from contaminated fields, 2005. *Epidemiol Infect.* 2008;136:157–65.
144. Klontz KC, Klontz JC, Mody RK, Hoekstra RM. Analysis of tomato and Jalapeño and Serrano pepper imports into the United States from Mexico before and during a National Outbreak of *Salmonella* serotype Saintpaul Infections in 2008. *J Food Prot.* 2010;73:1967–74.
145. Li B, Vellidis G, Liu H, Jay-Russell M, Zhao S, Hu Z, et al. Diversity and antimicrobial resistance of *Salmonella enterica* isolates from surface water in Southeastern United States. *Appl Environ Microbiol.* 2014;80:6355–65.
146. Kim S. *Salmonella* serovars from foodborne and waterborne diseases in Korea, 1998-2007: total isolates decreasing versus rare serovars emerging. *J Korean Med Sci.* 2010;25:1693–9.
147. Leclerc V, Moury F, Noel V, Berta-Vanrullen I, Noel V, Cadel-Six S, et al. Le réseau *Salmonella*, un dispositif de surveillance des salmonelles sur la chaîne alimentaire : bilan 2015. *Bull Épidémiologique Santé Anim Aliment.* 77 Numéro spécial – Surveillance sanitaire des aliments:1–81.
148. Jones N. How to stop data centres from gobbling up the world's electricity. *Nature.* 2018;561:163–6.

RÉSUMÉ

En vue d'une affiliation à l'école doctorale agriculture, alimentation, biologie, environnement, santé (n°0581), le présent mémoire a été rédigé en vue de l'obtention de l'habilitation à diriger des recherches auprès de l'université Paris-Est, université dans laquelle j'ai obtenu mon doctorat et enseigné la microbiologie.

En application de la loi du 26 janvier 1984 sur l'enseignement supérieur et de l'arrêté du 23 novembre relatif à l'habilitation à diriger des recherches, cette dernière « *sanctionne la reconnaissance du haut niveau scientifique du candidat, du caractère original de sa démarche dans un domaine de la science, de son aptitude à maîtriser une stratégie de recherche dans un domaine scientifique ou technologique suffisamment large et de sa capacité à encadrer de jeunes chercheurs* ». D'après l'arrêté du 25 mai 2016 relatif à la formation doctorale, un chercheur ayant une habilitation à diriger des recherches a le droit de participer à la désignation du directeur de son l'école doctorale, examiner comme rapporteur des travaux de thèse de doctorat et être responsable de l'encadrement d'une ou plusieurs thèses de doctorat.

Les découvertes de la microscopie au XVII^{ème} siècles, les travaux en bactériologie et parasitologie au XIX^{ème} siècle, l'essor de la biologie moléculaire au XX^{ème} siècle, ainsi que la révolution des disciplines omiques au XXI^{ème}, ont permis à la microbiologie d'aider à résoudre une succession de crises sanitaires qui ont frappé l'humanité depuis le néolithique jusqu'à nos jours. La crise sanitaire contemporaine est fortement liée à la nécessité de nourrir 9,7 milliards d'humains d'ici 2050 selon l'organisation des nations unies. Pour cette raison, les gouvernances en santé publique nationale, continentale et internationale d'aujourd'hui se basent sur les technologies les plus avancées, notamment le séquençage de génomes complets, pour contrôler et prévoir les épidémies, ainsi qu'apporter des solutions de protection aux populations.

Docteur en science et technique et l'environnement de l'école des ponts ParisTech en 2011, j'ai réalisé ma thèse de doctorat à l'université Paris-Est de Créteil en collaboration avec les laboratoires Eau de Paris, l'assistance publique hôpitaux de Paris et l'institut polytechnique VirginiaTech aux États-Unis. Entre 2011 et 2014, j'ai mené ma recherche comme postdoctorant à l'institut de recherche du centre de santé de l'université de McGill au Canada. Depuis 2015, je mets en place la génomique bactérienne dans le cadre de la mission « génome analyse modélisation risque » du laboratoire de sécurité des aliments de l'agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail.

Mon environnement de travail orienté vers une numérisation globale des données de santé publique nécessaire à l'implémentation de la génomique par les acteurs concernés est en adéquation avec ma carrière personnelle orientée vers l'amélioration continue des pratiques analytiques dans les différents domaines d'application de la microbiologie. Je propose donc dans ce mémoire un programme de recherche visant à apporter des solutions en sécurité alimentaire sur la base de mes travaux alliant la bactériologie, la biologie moléculaire et la génomique bactérienne.

Dans le contexte actuel d'une crise sanitaire de l'alimentation et d'une révolution technologique dans le domaine du séquençage, l'obtention d'une habilitation à diriger des recherches me permettra de poursuivre l'implémentation de la génomique bactérienne en sécurité alimentaire via des collaborations nationale et internationale, des recherches de financements, des encadrements de thèse de doctorat, des valorisations de mes travaux et des animations de la recherche.