

Université de Lille1
-Sciences et Technologies-



Mémoire présenté à l'Université de Lille1 Sciences et Technologies pour
l'obtention de l'Habilitation à Diriger des Recherches
Ecole Doctorale Biologie Santé de Lille (EDBSL)

Régulation de la virulence chez *Bordetella*
pertussis, agent responsable de la coqueluche

Docteur David Hot

Professeur Patrick Berche Directeur général de l'Institut Pasteur de Lille, membre correspondant de l'Académie Nationale de Médecine.

Professeur Jean Marie Lacroix (rapporteur) Professeur des Universités en génétique et microbiologie. Université Lille 1 Unité de Glycobiologie Structurale et Fonctionnelle (UGSF UMR CNRS-Lille1 8576).

Docteur Camille Locht (rapporteur) Directeur de recherche INSERM, Directeur du Centre d'Infection et d'Immunité de Lille (CIIL), U1019, UMR8204, Université de Lille.

Professeur Benoit Guery (rapporteur) PU PH, Université de Lille - Faculté de Médecine de Lille - CHRU Lille / Maladies infectieuses. Recherche translationnelle relations hôte-pathogènes EA7366.

Professeur Philippe Delannoy Professeur des universités. Université Lille 1, Unité de Glycobiologie Structurale et Fonctionnelle (UGSF UMR CNRS-Lille1 8576).

Professeur Yves Lemoine (garant) Professeur des universités. Université de Lille 1. Transcriptomique et Génomique Appliquée ó Centre d'Infection et d'Immunité de Lille (U1019, UMR 8204).

Le 6 juin 2016

Laboratoire de Transcriptomique et Génomique Appliquée (TAG), Centre d'Infection et d'Immunité de Lille (CIIL), Institut Pasteur de Lille

Remerciements

Il ne nous est que trop rarement donné l'occasion de remercier les personnes qui partagent notre quotidien professionnel. Je voudrais profiter de cette pause et de ce bilan que représente l'HDR pour sincèrement remercier les personnes dont j'ai croisé la route et qui m'ont accompagné pendant quelques jours ou pour plusieurs années. Ils m'ont tous apporté énormément, et bien entendu le chemin n'aurait pas été le même sans ces rencontres.

Au premier plan de ces remerciements je voudrais mentionner mes collaborateurs proches du laboratoire de Transcriptomique et Génomique Appliquée (ex- Biopuces). Yves Lemoine, bien entendu, qui a été un guide depuis mes commencements à l'Institut Pasteur de Lille et qui m'a aidé à prendre confiance et à évoluer aussi bien dans les aspects scientifiques que dans mes responsabilités de gestion de projets et d'équipe. Les autres membres de l'équipe, présente et passée, Patricia Paul, Ludovic Huot, Stéphanie Slupek, Ségolène Caboche, Anca Lucau, Delphine Beury, Erwan Levillain, Renaud Blervaque avec qui je partage ou j'ai partagé de grands moments et un quotidien rempli autant de travail constructif que d'amitié. J'y associe également les étudiants en thèse Christine Hubans, Alexandre D'Halluin et Léa Siegwald. Les membres de l'équipe GDscan de Gènes Diffusion, Christophe Audebert, Gaël Even, Sophie Merlin, Alexandre Loywick, Ludivine Liétar, Romain Dassonneville qui sont venus renforcer nos liens d'équipe et d'amitié depuis de nombreuses années maintenant.

Je voudrais aussi remercier les membres de l'équipe de Camille Locht avec qui j'ai la chance d'échanger et de tisser des liens forts de travail et d'amitié autour d'une passion commune pour *Bordetella pertussis* : Camille Locht, Rudy Antoine, Françoise Jacob-Dubuisson, Dominique Raze, Anne-Sophie Debrie, Loïc Coutte, (j'en oublie certainement, veuillez m'en excuser).

J'ai une pensée particulière pour ma famille, mes enfants, mes parents, mes amis qui sont un soutien permanent, un rempart contre le doute et un réconfort dans les moments plus sombres.

Enfin j'exprime toute ma gratitude aux membres du jury qui me font l'honneur de juger mon mémoire et ma soutenance d'HDR. Je suis heureux de bénéficier de leur regard éclairé.

Curriculum Vitae

ETAT CIVIL

Nom : Hot

Prénom : David

Né : le 25 mai 1970 à Douai (Nord)

Nationalité : Française

Situation familiale : marié, 2 enfants

Adresse personnelle : 7, avenue du champ de courses – 59130 Lambersart

Téléphone : 03 20 93 19 75

SITUATION PROFESSIONNELLE

Chargé de Recherche (CR2, IPL) depuis 2000. Dirige le laboratoire 'Transcriptomics and Applied Genomics'.

En charge du fonctionnement de la plate-forme : collaborations et partenariats scientifiques, développement nouvelles approches de génomique, gestion de l'équipe et du laboratoire.

En charge de programme de recherche de l'équipe sur les mécanismes moléculaires associés aux petits ARN régulateurs chez *Bordetella pertussis*.

Le laboratoire 'Transcriptomics and Applied Genomics' (initialement 'Laboratoire des Biopuces') a été créé en mars 2000. En octobre 2001, le laboratoire a reçu le label « GenHomme ». Il a intégré la Génopole de Lille en juin 2002 et a été rattaché à l'UMR CNRS 8161 dirigé par Dr. Yvan De Launois puis à l'IFR 142 dirigé par Camille Locht en janvier 2006. Notre laboratoire a ensuite rejoint lors de sa création en 2006 le Centre d'Infection et d'Immunologie de Lille (CIIL) dirigé par Dr. Camille Locht.

Adresse professionnelle

Laboratoire 'Transcriptomics and Applied Genomics' (TAG)

Institut Pasteur de Lille

Bâtiment Guérin, 2^{ème} étage

1, rue du professeur Calmette

BP245 – 59019 LILLE Cedex

Téléphone : 03 20 87 72 09

Fax : 03 20 87 73 11

E-mail : david.hot@pasteur-lille.fr

DIPLÔMES

- Juin 1999 **Doctorat d'Université (Ph.D.)** - Institute for Animal Health - British Biotechnology Science Research Council (BBSRC) Compton. School of Animal and Microbial Sciences, Université de Reading (Berkshire, UK).
Directeur de thèse : Dr. M.A. Skinner
Sujet : Etude de l'ARN polymérase ARN dépendante du virus de Gumboro (Infectious bursal disease virus, IBDV).
Jury : Dr. S Heaphy, Université de Leicester (UK), Dr. J.A. Hiscox, Université de Reading.
- 1994 **Ingénieur en Biotechnologie** - Ecole Supérieure de Biotechnologie de Strasbourg (ESBS).
- 1994 **DEA de biologie cellulaire et moléculaire** - Université Louis Pasteur, Strasbourg (mention bien).
- 1991 **DEUG de Biochimie** - Université de Lille I (mention assez bien).
- 1988 **Baccalauréat D** - Lycée St Jean Douai (mention assez bien).

FONCTIONS

- 2015-Auj. **Chef d'équipe** de la plateforme Transcriptomique et Génomique Appliquée.
Centre d'Infection et d'Immunité de Lille (CIIL) - IPL, Inserm U1019, UMR8204, Univ-Lille Nord de France.
- 2010-2014 **Responsable** de la plateforme Transcriptomique et Génomique Appliquée.
Centre d'Infection et d'Immunité de Lille (CIIL) - IPL, Inserm U1019, UMR8204, Univ-Lille Nord de France depuis 2010.
- 2012-Auj. **Co-initiateur et co-responsable** du programme « Pegase Biosciences » : Programme joint Institut Pasteur de Lille - Gènes Diffusion.
- 2000-2010 **Chargé de recherches** – Plate-forme Biopuces (Professeur Yves Lemoine). Institut Pasteur de Lille.
Groupe labellisé dans l'UMR 8161.

- 1999-2000 **Post-doc** - Laboratoire de virologie - Département Microbiologie des Ecosystèmes (Dr. Annick Mercenier). Institut Pasteur de Lille.
- 1995 **ETCA (Etablissement Technique Central de l'Armement)** - Arcueil. Biologie moléculaire.
Scientifique du contingent : Evaluation du SOS-chromotest comme dosimètre d'effet.
- 1994 **SmithKline Beecham Biologicals** Rixensart (Belgique) – Biologie moléculaire, virologie.
Stagiaire (6 mois) : Expression du second cadre de lecture ouverte (ORF) du virus de l'hépatite E en lignée stable (cellule CHO-K1). Caractérisation du produit d'expression. Détection de 'virus-like particles'.
- 1993 **SmithKline Beecham Biologicals** Rixensart (Belgique) – Biologie moléculaire, virologie.
Stagiaire (2 mois) : Evaluation du système d'expression de 'Semliki forest virus' Expression des protéines gD et ICP27 de Herpes Simplex Virus.
- 1992 **Biozentrum**, Bâle (Suisse) – Biologie moléculaire.
Stagiaire (2 mois) : Purification et cristallisation d'une protéine membranaire (porine) de *E.coli*.

Activités d'enseignement - Jury - Animation scientifiques

Activités d'enseignement

Cours

Master Pro Génie Cellulaire et Moléculaire (GCM) module 'Marqueurs moléculaires. Nanotechnologies', 2h de cours 2006 à 2012, 6h de cours en 2013 et 2014.

Formation scientifique Pasteur

'Analyse Transcriptomique par la Technologie des Puces à ADN', 30h de cours annuellement, organisation et animation, de 2005 à 2012.

Journée thématique

JTB4 M2R Biologie Santé, Lille. Etude Transcriptomique et Génomique Appliquée, bioinformatique associée, 1h de cours annuellement, de 2004 à 2014.

DUER

Diplôme Universitaire d'Ethique de la Recherche Biomédicale, Université Catholique de Lille. 2h de cours, mai 2005, juin 2006, avril 2007.

Formation complémentaire professionnalisante (FCP)

Master pro 'Génie cellulaire et moléculaire', Lille1. 4,5h de TP et 2h de cours, janvier 2005.

Travaux d'Etude et de Recherche

Master I 'Biologie - Biotechnologie', Lille1. 4,5h de TP et 2h de cours, janvier 2005.

Formation complémentaire professionnalisante

Master pro Protéomique Lille1. 35 h de TP, organisation et animation, février 2003, février 2004.

Jury

Evaluateur

ANR Appel à Projet Générique. Projet HydroGen : Métagénomique comparative comme instrument de mesure pour la biodiversité. Application à l'étude de la vie dans les océans. Institut National de Recherche en Informatique et en Automatique (INRIA) Centre Rennes - Bretagne Atlantique. Mai 2014

Evaluateur

Programme scientifique DIM Pathologies infectieuses humaines : mécanisme, risques et impact sociétal. Ile de France. 'Role of miRNAs in the innate response of cystic fibrosis bronchial epithelial cells infected with Pseudomonas aeruginosa' Inserm U938-Mucoviscidose : physiopathologie et phénogénomique. Hôpital St-Antoine Mars 2014.

Evaluation d'article

Evaluation d'un article de recherche pour le journal 'PLoS One' (PONE-D-12-16676) publications de 'the Public Library of Science'. Juillet 2012.

Jury

Recrutement d'un enseignant-chercheur, chaire d'excellence MCF 1147. Campagne « au fil de l'eau », Université de Lille2. Novembre 2010.

Jury

Assesseur pour la Thèse de Docteur en Pharmacie de M. Patrick Caneiro, Faculté des Sciences Pharmaceutiques, Université de Lille2. Octobre 2009.

Evaluateur

Programme scientifique 'Microarrays to map pertussis adaptation' pour le Dutch National Institute of Public Health and the Environment (acronyme RIVM) section "Strategic Research RIVM (SOR)". Avril 2009

Membre de jury

Assesseur pour la Thèse de Docteur en Pharmacie de Jean-Charles Mraz, Faculté des Sciences Pharmaceutiques, Université de Lille2. Février 2009.

Membre de jury

Stage de L. Laborde (encadrant : Anca Lucau) validation M1 IUP Génomique et protéomique, USTL. Septembre 2007.

Membre de jury

Examineur pour la Thèse d'Université de Sylvain Legrand, USTL. Décembre 2006.

Membre de jury (interne IPL)

Attribution des bourses de thèse 'région'. Mars 2006.

Evaluation d'article

Evaluation d'un article de recherche pour le journal 'Research in Microbiology', publications de l'Institut Pasteur aux éditions Elsevier. Juillet 2005.

Animations scientifiques

Représentant de l'Institut Pasteur de Lille auprès du réseau 'Institut Pasteur International Network BIO-IT Initiative' (IPIN-BIO-IT) animé par Magnus Fontes.

Co-organisation et animation d'une journée symposium sur les 'autour de la Biologie Intégrative' dans le cadre du PPF Bioinformatique Lille. 17 décembre 2014, Lille.

Co-organisation et animation d'une journée symposium sur les 'La protéomique à l'ère du tout omique' dans le cadre du PPF Bioinformatique Lille et de la FRE3642. 3 février 2014, Lille.

Co-organisation et animation d'une journée symposium sur les 'Phylogénomique' dans le cadre du PPF Bioinformatique Lille. 3 avril 2013, Lille.

Co-organisation et animation d'une journée symposium sur les 'Analyse bio-informatique des données de méta-génomique' dans le cadre du PPF Bioinformatique Lille et de l'IFR142. 19 décembre 2012, Lille.

Co-organisation et animation d'une journée symposium sur les 'Analyse bio-informatique des données NGS' dans le cadre du PPF Bioinformatique Lille et de l'IFR142. 7 décembre 2011, Lille.

Co-organisation et animation d'une journée symposium sur les 'Fouille de texte pour la biologie' dans le cadre du PPF Bioinformatique Lille. 20 septembre 2011, Lille.

Co-organisation et animation d'une journée symposium sur les 'Intestinal Flora: from physiology to metagenomic analysis and vice versa' dans le cadre du CIIL. 17 mars 2011, Lille.

Co-organisation et animation d'une journée symposium sur les 'Intégration des données hétérogènes' dans le cadre du PPF Bioinformatique Lille, de l'IFR142 et de l'IFR114. 4 mai 2010, Lille.

Organisation et animation d'une journée symposium sur les 'Next Generation Sequencing' dans le cadre du PPF Bioinformatique Lille et de l'IFR142. 9 mars 2009, Lille.

Représentant de l'Institut Pasteur de Lille auprès du PPF bioinformatique de Lille animé par Hélène Touzet.

Membre du comité d'organisation des Journées Ouvertes de Biologie Informatiques Mathématiques 2008 (JOBIM 2008). 30 juin, 1-2 juillet 2008, Lille.

Encadrement scientifique

Encadrements de Stages

Alexandre D'Halluin

Directeur de stage Master 2 - Lille1. Identification et caractérisation de petits ARN non-codants chez *Bordetella pertussis*, implication dans la régulation de la virulence. Octobre 2014 – Juin 2015. Taux d'encadrement : 100%.

Lauralie Péronne

Directeur de stage Licence Lille1. Étude d'un petit ARN non codant BprC potentiellement impliqué dans la régulation de la virulence de *Bordetella pertussis*. Mai – juin 2013. Taux d'encadrement 100%.

Sara Chabi

Directeur de stage 1ère année Master Biologie et Biotechnologie de Lille1. Etude des départs de transcription par dRNA-Seq chez *Bordetella pertussis*. Mars 2013 – juillet 2013. Taux d'encadrement 100%.

Florent Tiberghien

Directeur de stage 2ème année BTS Biotechnologies. Lycée Valentine LABBE de La Madeleine. Recherche de petits ARN non-codants potentiellement impliqués dans la régulation de la virulence chez *Bordetella pertussis*. Mai 2013 – juillet 2013. Taux d'encadrement 100%.

Florent Tiberghien

Directeur de stage 1ère année BTS Biotechnologies. Lycée Valentine LABBE de La Madeleine. Prédiction de petits ARNs chez *Bordetella pertussis* par puces à ADN. Mai 2012 – juillet 2012. Taux d'encadrement 100%.

Adam Ayache

Directeur de stage 2ème année d'Ecole Centrale Lille. Solution d'analyse des données issues du séquençage de transcriptome. Juin 2010 – août 2010. Taux d'encadrement 100%.

Omar Douffi

Directeur de stage de Licence professionnelle 'Biotechnologies Option Système d'Information et de Modélisations Appliqués à la Bioinformatique', Université d'Auvergne. *Exploration des outils d'analyse des technologies de séquençage à très haut débit « Next Generation Sequencing »*. Février 2009 – juin 2009. Taux d'encadrement : 100%.

Ludovic Debure

Directeur de 1ère année BTS Biotechnologies. Lycée Valentine LABBE de La Madeleine. Préparation et validation de puces à ADN. Mai 2008 – juillet 2008. Taux d'encadrement 100%.

2ème année BTS Biotechnologies. Lycée Valentine LABBE de La Madeleine. Transcriptomique génomique bactérienne. Janvier 2009 – mars 2009. Taux d'encadrement 100%.

Marion Réocreux

Directeur de stage 1ère année Master Génomique et Protéomique IUP de Lille1. Contribution à l'identification du rôle physiologique d'un nouveau petit ARN non-codant chez *Bordetella pertussis*. Mars 2008 – juillet 2008. Taux d'encadrement 100%.

Céline Duval

Directeur de stage de 2ème année d'IUP génomique et protéomique, Lille1. *Validation biologique des prédictions de petits ARN non-codants de chez *Bordetella pertussis**. Mars 2007 – août 2007. Taux d'encadrement : 100%.

Antra Elmousselly

Directeur de stage de DEA en statistique et informatique appliquées, Gembloux (Belgique). *Comparaison de différentes méthodes d'analyse statistique visant à mettre en*

évidence les gènes différentiellement exprimés dans les expérimentations basées sur l'utilisation des biopuces. Février 2007 – mai 2007. Taux d'encadrement : 50%.

Pascal Basselet

Directeur de stage de 2^{ème} année d'IUP génomique et protéomique, Lille1. *Etude de faisabilité de la détection de pathogènes alimentaires par biopuce génomique.* Février 2006 – septembre 2006. Taux d'encadrement : 50%.

Anthony D'Hondt

Directeur de stage de Master Pro Bioinformatique, Lille1. *Recherche d'ARN non-codants chez Bordetella pertussis.* Mars 2006 - septembre 2006. Taux d'encadrement : 100%.

Bérénice Wulbrecht

Directeur de stage de Master Pro Bioinformatique, Rouen. *Etude des données génomiques et transcriptomiques des Bordetelles dans le cadre de la virulence de la bactérie.* Septembre 2005 - juillet 2007. Taux d'encadrement : 50%.

Even Gaël

Directeur de stage de Master Pro Bioinformatique, Lille1. *Recherche de reporters pour l'élaboration d'une puce à ADN. Développement d'outils bioinformatiques pour les données issues d'études transcriptomiques.* Mars-octobre 2005. Taux d'encadrement : 100%.

Bérénice Wulbrecht

Directeur de stage de 3^{ème} année d'IUP génomique et protéomique, Lille1. *Mise en place d'une puce de détection et de typage des Papillomavirus Humains muqueux.* Mars-août 2005. Taux d'encadrement : 100%.

Pamelard Fabien

Directeur de stage de Master Pro Bioinformatique, Rouen. *Etude de faisabilité du typage des HPV par puce à ADN. Installation et optimisation de BASE BioArray Software Environment.* Septembre 2003 - juillet 2005. Taux d'encadrement : 50%.

Vermersch Mélanie

Directeur de stage de Licence professionnelle 'Biologie Appliquée à l'Entreprise', Orsay. *Etude des variations liées aux facteurs expérimentaux d'une puce à ADN représentative du génome de Bordetella pertussis.* Avril-juillet 2004. Taux d'encadrement : 100%.

Bayard Mathilde

Directeur de stage de Maîtrise, Lille1. *Etude de la correspondance entre un profil de reporters obtenu sur Bioanalyzeur Agilent / spectrométrie et son hybridation.* Juin-juillet 2003. Taux d'encadrement : 100%.

Bérénice Wulbrecht

Directeur de stage de 1^{ère} année IUP génomique et protéomique, Lille1. *Mise au point des conditions de PCR en vue de l'élaboration d'une puce Bordetella pertussis.* Septembre - novembre 2003. Taux d'encadrement : 100%.

Winogrodzki Sandrine

Directeur de stage de Master de protéomique, Lille 1 (majore de promotion). *Mise au point d'une puce à ADN pour la détection de pathogènes dans la semence porcine.* Mars-octobre 2003. Taux d'encadrement : 100%.

Petit Philippe

Directeur de stage de Master Pro de protéomique, Lille 1. *Approche d'une technique de génotypage sur biopuce du Mycobacterium basée sur l'utilisation de minisatellites.* Mars-octobre 2003. Taux d'encadrement : 100%.

Hasna Mehdi

Directeur de stage de Master Pro de protéomique, Lille 1. *Exploration des méthodologies de génotypage sur lame de verre.* Mars-octobre 2002. Taux d'encadrement : 100%.

Lorte Vanessa

Directeur de stage de 1^{ère} année Ecole Nationale Supérieure de Chimie de Lille, Lille1. *Optimisation des conditions de PCR sur une banque de clones d'ADN humains.* Juin-juillet 2002. Taux d'encadrement : 100%.

Cottineau Julien

Directeur de stage de 2^{ème} année BTS Biotechnologie Valentine Labbé. *Initiation à la technologie des puces à ADN.* Juin 2001. Taux d'encadrement : 100%.

Encadrement de Thèse d'Université

Léa Siegwald

Co-encadrement de thèse d'Université de l'école Doctorale Biologie et Santé Lille1/Lille2 en contrat CIFRE. Développement d'un pipeline intégrant des approches d'analyse innovantes et des statistiques multivariées pour la métagénomique par séquençage haut-débit. Début de thèse : mars 2014. Taux d'encadrement 25%.

Christine Hubans

Co-encadrement de thèse d'Université de l'école Doctorale Biologie et Santé de Lille1/Lille2. *Méthode ab initio de prédiction d'opérons chez les procaryotes et validations*

biologiques chez les Bordetelles. Soutenue le 27 octobre 2006. Taux d'encadrement 25%.

Liste des travaux

Publies écrites	Rang de signature				Moyenne des facteurs d'impacts
	1er	2ème	Dernier	coauteur	
2000				1	4,95
2003	2	1		1	3,2
2004				2	3,43
2005				1	1,85
2006		1			3,13
2007		1		2	7,66
2008				1	3,29
2009		2		1	9,33
2010				2	4,45
2011	1			2	6,4
2012				7	4.11
2013				7	6,53
2014			2	3	4,28
2015	1			3	3,97
Total = 44	4	5	2	33	
Communications orales ou affichées	1er	2ème	Dernier	coauteur	
1997	1	1			
1998	1				
1999		1			
2002	1				
2003	1				
2004		2		1	
2005	1	2			
2006	3	1		1	
2007	2	1	1	2	
2008				1	
2010	2				
2011	2				
2013			1		
2014				1	
2015	1				
Total = 31	15	8	2	6	

Publications internationales dans des journaux à comité de lecture associées à l'activité de plateforme de génomique et de bioinformatique

Fougeron D, Van Maele L, Songhet P, Cayet D, **Hot D**, Van Rooijen N, Mollenkopf HJ, Hardt WD, Benecke AG, Sirard JC. Indirect Toll-like receptor 5-mediated activation of conventional dendritic cells promotes the mucosal adjuvant activity of flagellin in the respiratory tract. *Vaccine*. 2015 May 20. pii: S0264-410X(15)00654-4. doi: 10.1016/j.vaccine.2015.05.022. [Epub ahead of print] PubMed PMID: 26003491. IF2014=3.62

Spear R, Boytard L, Blervaque R, Chwastyniak M, **Hot D**, Vanhoutte J, Staels B, Lemoine Y, Lamblin N, Pruvot FR, Haulon S, Amouyel P, Pinet F. Adventitial Tertiary Lymphoid Organs as Potential Source of MicroRNA Biomarkers for Abdominal Aortic Aneurysm. *Int J Mol Sci*. 2015 May 18;16(5):11276-93. doi:10.3390/ijms160511276. PubMed PMID: 25993295. IF2014=2.86

Reinhard A, Bressenot A, Dassonneville R, Loywick A, **Hot D**, Audebert C, Marchal S, Guillemain F, Chamaillard M, Peyrin-Biroulet L, Bezdetsnaya L. Photodynamic Therapy Relieves Colitis and Prevents Colitis-associated Carcinogenesis in Mice. *Inflamm Bowel Dis*. 2015 May;21(5):985-95. doi: 10.1097/MIB.0000000000000354. PubMed PMID: 25806846. IF2014=4.46

Audebert C, **Hot D**, Lemoine Y et Caboche S. Le séquençage haut-débit : vers un diagnostic basé sur la séquence complète du génome de l'agent infectieux, *Médecine/Sciences* 2014, 30 : 1144-1151. IF=NA

Caboche S, Audebert C, Lemoine Y, **Hot D**. Comparison of mapping algorithms used in high-throughput sequencing: application to Ion Torrent data. *BMC Genomics*. 2014 Apr 5;15(1):264. doi: 10.1186/1471-2164-15-264. PubMed PMID: 24708189; PubMed Central PMCID: PMC4051166. IF2014=3.99

Caboche S, Audebert C, **Hot D**. High-throughput sequencing, a versatile weapon to support genome-based diagnosis in infectious diseases: applications to clinical bacteriology. *Pathogens*. 2014, 3(2), 258-279; doi:10.3390/pathogens3020258. IF2015=NA

Mutez E, Nkiliza A, Belarbi K, de Broucker A, Vanbesien-Mailliot C, Bleuse S, Duflot A, Comptdaer T, Semaille P, Blervaque R, **Hot D**, Leprêtre F, Figeac M, Destée A, Chartier Harlin MC. Involvement of the immune system, endocytosis and EIF2 signaling in both genetically determined and sporadic forms of Parkinson's disease. *Neurobiol Dis*. 2014 Mar;63:165-70. doi: 10.1016/j.nbd.2013.11.007. Epub 2013 Nov 20. PubMed PMID: 24269915. IF2013 = 5.62

Van Maele L, Fougeron D, Janot L, Didierlaurent A, Cayet D, Tabareau J, Rumbo M, Corvo-Chamaillard S, Boulenouar S, Jeffs S, Vande Walle L, Lamkanfi M, Lemoine Y, Erard F, **Hot D**, Hussell T, Ryffel B, Benecke AG, Sirard JC. Airway structural cells regulate TLR5-mediated mucosal adjuvant activity. *Mucosal Immunol*. 2013 Sep 25. doi: 10.1038/mi.2013.66. [Epub ahead of print] PubMed PMID: 24064672. IF2011 = 7.00

Walker R, Gissot M, Huot L, Alayi TD, **Hot D**, Marot G, Schaeffer-Reiss C, Van Dorselaer A, Kim K, Tomavo S. Toxoplasma Transcription Factor TgAP2XI-5 Regulates the Expression of Genes Involved in Parasite Virulence and Host Invasion. *J Biol Chem*. 2013 Oct 25;288(43):31127-38. doi: 10.1074/jbc.M113.486589. Epub 2013 Sep 10. PubMed PMID: 24025328. IF2012 = 4.61

Tian TV, Tomavo N, Huot L, Flourens A, Bonnelye E, Flajollet S, **Hot D**, Leroy X, de Launoit Y, Duterque-Coquillaud M. Identification of novel TMPRSS2:ERG mechanisms in prostate cancer metastasis: involvement of MMP9 and PLXNA2. *Oncogene*. 2013 May 27. doi: 10.1038/onc.2013.176. [Epub ahead of print] PubMed PMID: 23708657. IF2012 = 7.35

Gissot M, Walker R, Delhaye S, Alayi TD, Huot L, **Hot D**, Callebaut I, Schaeffer-Reiss C, Dorselaer AV, Tomavo S. Toxoplasma gondii Alba Proteins Are Involved in Translational Control of Gene Expression. *J Mol Biol*. 2013 Feb 21. doi:pri: S0022-2836(13)00103-4. 10.1016/j.jmb.2013.01.039. [Epub ahead of print] PubMed PMID: 23454356. IF2011 = 4.00

Couturier-Maillard A, Secher T, Rehman A, Normand S, De Arcangelis A, Haesler R, Huot L, Grandjean T, Bressenot A, Delanoye-Crespin A, Gaillot O, Schreiber S, Lemoine Y, Ryffel B, **Hot D**, Núñez G, Chen G, Rosenstiel P, Chamaillard M. NOD2-mediated dysbiosis predisposes mice to transmissible colitis and colorectal cancer. *J Clin Invest*. 2013 Jan 2. doi:pri: 62236. 10.1172/JCI62236. [Epub ahead of print] PubMed PMID: 23281400. IF2011 =13.07

Fréalte E, Aliouat-Denis CM, Delhaes L, **Hot D**, Dei-Cas E. Transcriptomic Insights into the Oxidative Response of Stress-Exposed *Aspergillus Fumigatus*. *Curr Pharm Des*. 2012 Dec 26. [Epub ahead of print] PubMed PMID: 23278536. IF2011 = 3.87

Walker R, Gissot M, Croken MM, Huot L, **Hot D**, Kim K, Tomavo S. The Toxoplasma nuclear factor TgAP2XI-4 controls bradyzoite gene expression and cyst formation. *Mol Microbiol*. 2012 Dec 14. doi: 10.1111/mmi.12121. [Epub ahead of print] PubMed PMID: 23240624. IF2011 = 5.01

Flajollet S, Tian TV, Huot L, Tomavo N, Flourens A, Holder-Espinasse M, Le Jeune M, Dumont P, **Hot D**, Mallein-Gerin F, Duterque-Coquillaud M. Increased adipogenesis in cultured embryonic chondrocytes and in adult bone marrow of dominant negative *erg* transgenic mice. *PLoS One*. 2012; 7(11):e48656. doi: 10.1371/journal.pone.0048656. Epub 2012 Nov 14. PubMed PMID: 23155398; PubMed Central PMCID: PMC3498236. IF2011 = 4.09

Lucau-Danila A, Toitot C, Goulas E, Blervacq AS, **Hot D**, Bahrman N, Sellier H, Lejeune-Hénaut I, Delbreil B. Transcriptome analysis in pea allows to distinguish chilling and acclimation mechanisms. *Plant Physiol Biochem*. 2012 Sep; 58:236-44. doi: 10.1016/j.plaphy.2012.07.012. Epub 2012 Jul 20. PubMed PMID: 22858528. IF2011 = 2.84

Elass-Rochard E, Rombouts Y, Coddeville B, Maes E, Blervaque R, **Hot D**, Kremer L, Guérardel Y. Structural determination and Toll-like receptor 2-dependent proinflammatory activity of dimycolyl-diarabino-glycerol from *Mycobacterium marinum*. *J Biol Chem*. 2012 Oct 5; 287(41):34432-44. doi: 10.1074/jbc.M112.378083. Epub 2012 Jul 13. PubMed PMID: 22798072; PubMed Central PMCID: PMC3464548. IF2011 = 4.77

Bougault C, Aubert-Foucher E, Paumier A, Perrier-Groult E, Huot L, **Hot D**, Duterque-Coquillaud M, Mallein-Gerin F. Dynamic compression of chondrocyte-agarose constructs reveals new candidate mechanosensitive genes. *PLoS One*. 2012; 7(5):e36964. Epub 2012 May 17. PubMed PMID: 22615857; PubMed Central PMCID: PMC3355169. IF2011 = 4.09

Gissot M, Walker R, Delhaye S, Huot L, **Hot D**, Tomavo S. *Toxoplasma gondii* chromodomain protein 1 binds to heterochromatin and colocalises with centromeres and telomeres at the nuclear periphery. *PLoS One*. 2012; 7(3):e32671. Epub 2012 Mar 9. PubMed PMID: 22427862; PubMed Central PMCID: PMC3302879. IF2011 = 4.09

Normand S, Delanoye-Crespin A, Bressenot A, Huot L, Grandjean T, Peyrin-Biroulet L, Lemoine Y, **Hot D**, Chamaillard M. Nod-like receptor pyrin domain-containing protein 6 (NLRP6) controls epithelial self-renewal and colorectal carcinogenesis upon injury. *Proc Natl Acad Sci U S A*. 2011 Jun 7; 108(23):9601-6. Epub 2011 May 18. PubMed PMID: 21593405; PubMed Central PMCID: PMC3111299. IF2011 = 9.68

Pierrot C, Acroute Dit Vampouille A, Vandomme A, Lafitte S, Pierce RJ, **Hot D**, Khalife J. Gene profiling analysis reveals the contribution of CD24 and P2Y6R to the susceptibility of young rats to *Plasmodium berghei* infection. *Cell Microbiol*. 2011

May;13(5):752-63. doi:10.1111/j.1462-5822.2011.01574.x. [Epub ahead of print]
PubMed PMID: 21323829. IF2011 = 5.46

Lucau-Danila A, Laborde L, Legrand S, Huot L, **Hot D**, Lemoine Y, Hilbert JL, Hawkins S, Quillet MC, Hendriks T, Blervacq AS. Identification of novel genes potentially involved in somatic embryogenesis in chicory (*Cichorium intybus* L.). *BMC Plant Biol.* 2010 Jun 22; 10:122. PubMed PMID: 20565992. IF2011 = 3.45

Lamblin N, Ratajczak P, **Hot D**, Dubois E, Chwastyniak M, Beseme O, Drobecq H, Lemoine Y, Koussa M, Amouyel P, Pinet F. Profile of macrophages in human abdominal aortic aneurysms: a transcriptomic, proteomic, and antibody protein array study. *J Proteome Res.* 2010 Jul 2;9(7):3720-9. PubMed PMID: 20513153. IF2011 = 5.11

Chapuis J., **Hot D.**, Hansmannel F., Kerdraon O., Ferreira S., Hubans C., Maurage C.A., Huot L., Bensemain F., Laumet G., Ayrat A.M., Fievet N., Hauw J.J., Dekosky S.T., Lemoine Y., Iwatsubo T., Wavrant-Devrièze F., Dartigues J.F., Tzourio C., Buée L., Pasquier F., Berr C., Mann D., Lendon C., Alperovitch A., Kamboh M.I., Amouyel P., Lambert J.C. Transcriptomic and genetic studies identify IL-33 as a candidate gene for Alzheimer's disease. *Mol Psychiatry.* 2009 Feb 10. [Epub ahead of print]

Bensemain, F., **Hot, D.**, Ferreira, S., Dumont, J., Bombois, S., Maurage, CA., Huot, L., Hermant, X., Levillain, E., Hubans, C., Hansmannel, F., Chapuis, J., Hauw, JJ., Schraen, S., Lemoine, Y., Buée, L., Berr, C., Mann, D., Pasquier, F., Amouyel, P., Lambert, JC. (2007). Evidence for induction of the ornithine transcarbamylase expression in Alzheimer's disease. *Mol Psychiatry.* 2009 Jan;14(1):106-16. Epub 2007 Sep 25. Facteur d'impact (FI) = 11,804.

Pierrot, C., Adam, E., **Hot, D.**, Lafitte, S., George, J., Capron, M., George, J., Khalife, J. (2007). Contribution of T cells and Neutrophils in protection of young susceptible rats from fatal experimental malaria. *J of Immunology* **178**(3), 1713-1722. FI = 6,293.

Jeshtadi, A., Henriquet, G., Laidlaw, S. M., **Hot, D.**, Zhang, Y., and Skinner, M. A. (2005). In vitro expression and analysis of secreted fowlpox virus CC chemokine-like proteins Fpv060, Fpv061, Fpv116 and Fpv121. *Arch Virol* **150**(9), 1745-62. FI = 1,850.

Lefebvre, C., Cocquerelle, C., Vandenbulcke, F., **Hot, D.**, Huot, L., Lemoine, Y., and Salzet, M. (2004). Transcriptomic analysis in the leech *Theromyzon tessulatum*: involvement of cystatin B in innate immunity. *Biochem J* **380**(Pt 3), 617-25. FI = 4,100.

Hot, D., Legeay, O., Jacques, J., Gantzer, C., Caudrelier, Y., Guyard, K., Lange, M., and Andreoletti, L. (2003). Detection of somatic phages, infectious enteroviruses

and enterovirus genomes as indicators of human enteric viral pollution in surface water. *Water Res* **37**(19), 4703-10. FI = 2,459.

Lambert, J. C., Testa, E., Cognat, V., Soula, J., **Hot, D.**, Lemoine, Y., Gaypay, G., and Amouyel, P. (2003). Relevance and limitations of public databases for microarray design: a critical approach to gene predictions. *Pharmacogenomics J* **3**(4), 235-41. FI = 3,957.

Olivier, C., **Hot, D.**, Huot, L., Ollivier, N., El-Mahdi, O., Gouyette, C., Huynh-Dinh, T., Gras-Masse, H., Lemoine, Y., and Melnyk, O. (2003). Alpha-oxo semicarbazone peptide or oligodeoxynucleotide microarrays. *Bioconjug Chem* **14**(2), 430-9. FI = 3,823.

Andreoletti, L., Bourlet, T., Moukassa, D., Rey, L., **Hot, D.**, Li, Y., Lambert, V., Gosselin, B., Mosnier, J. F., Stankowiak, C., and Wattré, P. (2000). Enteroviruses can persist with or without active viral replication in cardiac tissue of patients with end-stage ischemic or dilated cardiomyopathy. *J Inf Dis* **182**(4), 1222-7. FI = 4,953.

Publications internationales dans des journaux à comité de lecture associées à l'activité de recherche sur *Bordetella pertussis*

Hot D*, Bibova I*, Keidel K, Amman F, Slupek S, Cerny O, Gross R, Vecerek B. Transcriptional profiling of *Bordetella pertussis* reveals requirement of RNA chaperone Hfq for Type III secretion system functionality. *RNA Biol.* 2015;12(2):175-85. [Epub ahead of print] PubMed PMID: 25674816.0. IF2014=4.97

*: co-first authors

Hoo R, Lam JH, Huot L, Pant A, Li R, **Hot D**, Alonso S. Evidence for a role of the polysaccharide capsule transport proteins in pertussis pathogenesis. *PLoS One.* 2014 Dec 12;9(12):e115243. doi: 10.1371/journal.pone.0115243. eCollection 2014. PubMed PMID: 25501560; PubMed Central PMCID: PMC4264864. IF2014=3.23

Bibova I, Skopova K, Masin J, Cerny O, **Hot D**, Sebo P, Vecerek B. The RNA chaperone Hfq is required for virulence of *Bordetella pertussis*. *Infect Immun.* 2013 Nov;81(11):4081-90. doi: 10.1128/IAI.00345-13. Epub 2013 Aug 26. Erratum in: *Infect Immun.* 2014 Jul;82(7):3087. PubMed PMID: 23980112; PubMed Central PMCID: PMC3811842.. IF2012 = 4.07

Hot D, Slupek S, Wulbrecht B, D'Hondt A, Hubans C, Antoine R, Loch C, Lemoine Y. Detection of small RNAs in *Bordetella pertussis* and identification of a novel

repeated genetic element. *BMC Genomics*. 2011 Apr 27;12:207. PubMed PMID: 21524285; PubMed Central PMCID: PMC3110155. IF2011 = 4.07

Lechner M, Schmitt K, Bauer S, **Hot D**, Hubans C, Levillain E, Locht C, Lemoine Y, Gross R. Genomic island excisions in *Bordetella petrii*. *BMC Microbiol*. 2009 Jul 18;9:141. PubMed PMID: 19615092; PubMed Central PMCID: PMC2717098. IF2011 = 3.04

Feunou, PF., Ismaili, J., Debrie, AS., Huot, L., **Hot, D.**, Raze, D., Lemoine, Y., Locht, C. Genetic stability of the live attenuated *Bordetella pertussis* vaccine candidate BPZE1. *Vaccine*. 2008 Oct 23;26(45):5722-7. Epub 2008 Aug 30.

Herrou, J., Bompard, C., Antoine, R., Leroy, A., Rucktooa, P., **Hot, D.**, Huvent, I., Locht, C., Villeret, V., Jacob-Dubuisson, F. (2007). Structure-based mechanism of ligand binding for periplasmic solute-binding protein of the Bug family. *J Mol Biol*. **373**(4), 954-64. FI = 4,890.

Caro, V., **Hot, D.**, Guigon, G., Hubans, C., Arrivé, M., Soubigou, M., Renauld-Mongénie, G., Antoine, R., Locht, C., Lemoine, Y and Guiso, N. (2006). Temporal analysis of French *Bordetella pertussis* isolates by comparative whole-genome hybridisation. *Microbes and Infection* **8**(8), 2228-35. FI = 3,127.

Locht, C., Antoine, R., Raze, D., Mielcarek, N., **Hot, D.**, Lemoine, Y., and Mascart, F. (2004). *Bordetella pertussis* from functional genomics to intranasal vaccination. *Int J Med Microbiol* **293**(7-8), 583-8. FI = 2,760.

Hot, D., Antoine, R., Renauld-Mongénie, G., Caro, V., Hennuy, B., Levillain, E., Huot, L., Wittmann, G., Poncet, D., Jacob-Dubuisson, F., Guyard, C., Rimlinger, F., Aujame, L., Godfroid, E., Guiso, N., Quentin-Millet, M. J., Lemoine, Y., and Locht, C. (2003). Differential modulation of *Bordetella pertussis* virulence genes as evidenced by DNA microarray analysis. *Mol Genet Genomics* **269**(4), 475-86. FI = 2,552.

Publication dans des ouvrages

Melnyk, O., **Hot, D.**, Olivier, C., Ollivier, N., Huot, L., Lemoine, Y., Gouyette, C.. Current Protocols in Nucleic Acid Chemistry. Chapter 12-06: Nucleic acid-based microarrays and nanostructures. Preparation of Alpha-Oxo Semicarbazone Oligonucleotide Microarrays.

Conférences par invitation

Hot D, Caboche S. A pipeline to quickly and automatically characterize microbial genomes from high-throughput sequencing data. INDA Steering Committee - Scientific Session 2015 #INDA4IP15, Institut Pasteur, Pasteur, 17 mars 2015 (https://www.youtube.com/watch?v=_kLPswallIN4&noredirect=1).

Gissot M, Walker R, Delhaye S, **Hot D**, Schaeffer-Reiss C, Van Dorsselaer A and Tomavo S. TgAlba proteins participate to the stress response through transcriptional and post-transcriptional regulation of gene expression. 11th International meeting on toxoplasmosis,

Huot L, Slupek S, Keidel K, Lemoine Y, **Hot D**. Characterisation of BprC an Antisense Small RNA of the 5'UTR of Bordetella pertussis Major Virulence Regulator Gene, bvgA. 10th International Symposium on Bordetella. Trinity Biomedical Sciences Institute. Trinity College Dublin, Ireland, 8-11 September 2013.

Hot, D. Insight into the world of Bordetellae small RNAs and their regulation. Academy of Sciences of the Czech Republic, Institute of Microbiology. Conférence "Bordetella pertussis virulence and vaccines: The battle continues ...". Prague (République Tchèque), Septembre 2011.

Fréalles E, Caneiro P, Ferreira S, Aliouat CM, Chabé M, **Hot D**, Dei-Cas E, Delhaes L. Microarrays and deep RNA sequencing for transcriptomic analysis of *Aspergillus fumigatus* hyphae exposed to oxidative stress. Valence (Espagne), octobre 2011.

Hot, D. Detection of small RNAs in *Bordetella pertussis*. Academy of Sciences of the Czech Republic, Institute of Microbiology. Conférence sur invitation du Dr. Peter Sebo, Laboratory of Molecular Biology of Bacterial Pathogens. Prague (République Tchèque), janvier 2011.

Hot, D. Detection of small RNAs in *Bordetella pertussis*. Eupertstrain / Eupertgenomics meeting. Bilthoven (Pays-Bas), juin 2010.

Hot, D. The use of arrays to study *Bordetella pertussis* virulence. Eupertstrain / Eupertgenomics meeting. Turku (Finlande), juin 2007.

Hot, D. Recherche de petits ARN non-codants chez *Bordetella pertussis*. Symposium IFR 142 « Petits ARNs et ARNs non-codants ». Institut Pasteur de Lille, mai 2007.

Hot, D. Nouveaux développements des techniques transcriptomiques. Club Francophone des Cellules Dendritiques (CFCD). Institut Gustave Roussy, Villejuif, décembre 2006.

Hot, D., Hubans, C. Environnement BASE, interconnexion Bioconductor/R. Journées RNG (Réseau National des Génopoles) Transcriptome et Bioinformatique. Toulouse, mars 2005.

Hot, D. Normalisation et standardisation des données biopuces bicanal/monocanal. Groupe de travail « Transcriptome et Régulation Transcriptionnelle » de l'Institut de Recherche Interdisciplinaire (IRI). Villeneuve D'Ascq, décembre 2003.

Communications orales en congrès scientifiques internationaux

Hot, D. The use of arrays to study *Bordetella pertussis* virulence. Eupertstrain/Eupertgenomics meeting. Turku (Finlande), juin 2007.

Antoine, R., **Hot, D.,** Renaud-Mongénie G., Levillain, E., Huot, L., Wittman, G., Arnaud-Barbe, F., Quentin-Millet, M.J., Aujame, L., Lemoine, Y., Loch, C. The genome sequence of *Bordetella pertussis* 18323. Microbial Genomes 2007, Wellcome Trust Conference center. Hinxton, UK, avril 2007.

Hubans, C., **Hot, D.,** Lemoine, Y. and Mephu-Nguifo, E. Hybrid approach based on a combination of methods for operon prediction. European Conference on Computational Biology (ECCB), student symposium. Madrid (Espagne), septembre 2005.

Hubans, C., **Hot, D.,** Lemoine, Y. and Mephu-Nguifo, E. An algorithm for prediction of operon structures in prokaryotes. Journée Ouvertes Biologie Informatique Mathématiques (JOBIM). Montréal (Canada), juillet 2004.

Andréoletti, L., **Hot, D.,** Debrie, A.S., Lesay, M., Wattré, P., Lange, M., Mercenier, A. Rapid detection of infectious enteroviruses in environmental samples by an integrated cell culture-PCR assay. XIIth International Congress of Virology. Sydney, Australie, août 1999.

Hot, D. and Skinner, M. Expression and preliminary characterisation of the RNA dependent RNA polymerase (VP1) of Infectious Bursal Disease Virus (IBDV). VIth International symposium on double stranded RNA viruses. Mexico, Mexique, novembre 1997

Communications affichées en congrès scientifiques internationaux

Walker R, Huot L, **Hot D**, Tomavo S, Gissot M. Toxoplasma gondii sequence-specific transcription factor AP2XI-4 is involved in bradyzoite stage conversion. Molecular Parasitology Meeting, Woods Hole (USA), septembre 2014.

Caboche S, Audebert C, Even G, Blervaque R, Lemoine Y, **Hot D**. Pipeline for automatic analysis of Whole Genome Sequencing data: Application to bacterial genome characterization. Genomes 2014: EMBO Conference on Microbiology after the genomics revolution, Institut Pasteur, Paris, juin 2014.

Hot D, Huot L, Slupek S, Keidel K, Lemoine Y. Characterization of BprC an antisense of the 5'UTR of Bordetella pertussis major virulence regulator, BvgA. ASM conference on "Regulating with RNA in Bacteria". Würzburg (Allemagne), juin 2013.

Hot, D., Slupek, S., Wulbrecht, B., D'Hondt, A., Hubans, C., Antoine, R., Locht, C., Lemoine, Y. Detection of small RNAs in Bordetella pertussis. ASM conference on "Regulating with RNA in Bacteria". San Juan (Porto Rico), Mars 2011.

Hot, D., Slupek, S., Wulbrecht, B., D'Hondt, A., Hubans, C., Antoine, R., Locht, C., Lemoine, Y. Detection of small RNAs in Bordetella pertussis. 9th International Bordetella Symposium. Baltimore (USA), Octobre 2010.

Lucau-Danila, A., Laborde, L., Huot, L., Couillerot, J-P., **Hot, D.**, Hilbert, J-L., Blervacq, A-S. Microarray analyses distinguish genes involved in cell reactivation from genes involved in morphogenesis patterns during in vitro induction of Cichorium leaves. Journées RFP, BousSENS, France. 2008.

Wulbrecht, B., Hubans, C., **Hot, D.**, Lemoine, Y. Improving Operon prediction by DNA microarray data. Journées Ouvertes de Biologie Informatiques Mathématiques (JOBIM), juillet 2007.

Peyrin-Biroulet, L., Parmentier, E., Merour, E., Dubuquoy, L., Vignal, C., Erdual, E., **Hot, D.**, Levillain, E., Neut, C., Chavatte, P., Colombel, JF, Desreumaux, P., Chamailard, M. Peroxisome proliferator-activated receptor gamma functions as an antibacterial factor. 13th International Congress of Mucosal Immunology (ICMI). Shinagawa (Japon), juillet 2007.

Le Jeune, M., Huot, L., Dumont, P., Marchand, N., Tomavo, N., Flourens, A., Tian, T., Holder-Espinasse, M., **Hot, D.** et Duterque-Coquillaud, M. Conséquences transcriptionnelles de l'expression d'une protéine Ets à effet trans-dominant négatif

dans le cartilage embryonnaire. 10^{ème} Journées Françaises de Biologie des Tissus Minéralisés (JFBTM). St Valery sur Somme, mai 2007.

Hot, D., D'Hondt, A., Hubans, C., Levillain, E., Huot, L., Lemoine, Y. In silico search for small untranslated RNAs (sRNAs) in *Bordetella pertussis* genome. 8th International Symposium. Paris, Novembre 2006.

Hubans, C., **Hot, D.**, Huot, L., Wulbrecht, B., Slupek, S., Mephu-Nguifo, E. and Lemoine, Y. Operon prediction on *Bordetella* species using 'ab initio' approach of discrimination between transcribed and untranscribed sequences. 8th International Symposium. Paris, novembre 2006.

Herrou, J., Leroy, A., Bompard, C., Rucktooa, D., **Hot, D.**, Levillain, E., Loch, C., Antoine, R., Jacob-Dubuisson, F. Binding of nicotinic acid, a negative modulator of virulence in *Bordetella pertussis*, by the periplasmic protein Bug27. Paris, novembre 2006.

Hot, D., D'Hondt, A., Hubans, C., Levillain, E., Huot, L., Lemoine, Y. search for small untranslated RNAs in *Bordetella pertussis* genome using computational approach. Lille-Ghent 'Computational Biology' meeting. Villeneuve D'Ascq (France), juin 2006.

Legrand, S., Quillet, M.C., Lucau-Danila, A., **Hot, D.**, Huot, L., Lemoine, Y., Hilbert, J.L. Combining SSH, cDNA microarray and in silico analysis for identification of differentially expressed genes during somatic embryogenesis in *Cichorium intybus*. 6^{ème} colloque National de la SFBV, Arcachon, France, 2005.

Hubans, C., **Hot, D.**, Lemoine, Y., Mephu-Nguifo, E. Hybrid approach based on a combination of methods for operon prediction. European Conference on Computational Biology (ECCB), student symposium. Madrid (Espagne), septembre 2005.

Hubans, C., **Hot, D.**, Lemoine, Y. and Mephu-Nguifo, E. Un algorithme de prediction de structures opéroniques chez les prokaryotes. Journée Ouvertes Biologie Informatique Mathématiques (JOBIM). Montréal (Canada), juillet 2004.

Soula, J., Levillain, E., Hubans, C., **Hot, D.** and Lemoine, Y. FindExpress : a software designed for the specific priming of prokaryotic mRNA. Journée Ouvertes Biologie Informatique Mathématiques (JOBIM). Montréal (Canada), juillet 2004.

Hot, D., Antoine, R., Renauld-Mongénie, G., Caro, V., Levillain, E., Huot, L., Jacob-Dubuisson, F., Guyard, C., Wittmann, G., Poncet, D., Aujame, L., Quentin-Millet, M.J., Lemoine, Y. and Loch, C. Differential expression (DE) of *Bordetella pertussis*

genes using cDNA microarrays. VIIth International symposium on pertussis: Genome, Pathogenesis and Immunity. Cambridge (U.K.), septembre 2002.

Hot, D., and Skinner, M. Expression and characterisation of the RNA dependent RNA polymerase (VP1) of Infectious Bursal Disease Virus (IBDV). International symposium on strategies in virus-host relationships. Lyon, février 1998.

Thomas, W., **Hot, D.** and Skinner, M. Alternative methods for reverse genetics of infectious bursal disease virus. VIth International symposium on double strand RNA viruses. Mexico, november 1997

HABILITATION A DIRIGER DES
RECHERCHES

DR. DAVID HOT

RAPPORT SCIENTIFIQUE

PREAMBULE

L'activité de mon équipe se partage en 2 axes principaux. Nous sommes plateforme technologique en génomique. Nous développons à ce titre des collaborations avec des laboratoires académiques du campus de l'Institut Pasteur de Lille et des universités/CHR de la région ou au national et à l'international avec d'autres instituts, universités ou industriels. Nos expertises concernent les technologies du séquençage à haut débit et les 'microarrays'. Nos applications récentes sont le séquençage et reséquençage de génomes microbiens entiers, le RNA-seq et 'differential RNA-seq' sur bactérie, la métagénomique sur l'ARNr 16S et la mesure d'expression de miRNA par puces à ADN. En parallèle sont développées des solutions de bioinformatique. Un pipeline d'analyse automatique de reséquençage de génomes bactériens a été mis en place. Il permet d'identifier les changements ponctuels (mutations et DIPs) ou plus importants (transfert de gènes, large délétion, acquisition de prophage et plasmide) entre un génome d'étude et un ou des génomes de références.

Nous sommes également laboratoire de recherche. Notre thématique de recherche porte sur l'étude de la régulation de virulence chez *Bordetella pertussis* (l'agent responsable de la coqueluche). C'est sur ce projet que porteront le manuscrit et la présentation de mon HDR. Ce projet a débuté au début des années 2000 par la mise au point et l'utilisation de la première puce à ADN pour ce pathogène. Cet outil nous a permis d'étudier précisément le régulon du régulateur principal de la virulence, BvgA. Nous avons ensuite comparé les caractéristiques génomiques de différentes souches de Bordetelles et isolats de *Bordetella pertussis* afin de mettre en relation les phénotypes de virulence et les caractéristiques géniques. Enfin, nous nous concentrons ces dernières années sur les voies de régulation moins classiques, notamment nous développons l'étude des petits ARNs non-codants et leurs potentielles implications dans la régulation des facteurs de virulence. Nous avons, tout d'abord, établi une première liste d'identification de petits ARNs chez les Bordetelles, puis nous avons étudié l'implication de ces petits ARNs non codants dans la virulence en caractérisant un mutant Hfq et en étudiant son comportement *in vitro* et *in vivo*. Un des petits ARNs, BprC, est antisens du 5'UTR de *bvgA* le gène codant pour le facteur de transcription majeur de la virulence. L'implication de BprC sur la régulation de la transcription ou de la traduction de *bvgA* ainsi que son mode d'action sont en cours d'étude. Enfin, une définition complète du transcriptome primaire incluant une liste exhaustive de petits ARNs non-codants chez *Bordetella pertussis* est actuellement finalisée par des approches de RNA-seq profond et de 'differential RNA-seq'. La liste des candidats de petits ARNs non-codants est en cours de validation et la fonction et mode d'action des plus intéressants seront étudiés dans le cadre de la régulation de la virulence.

CHAPITRE 1 : INTRODUCTION

1.1 *Bordetella pertussis* et la coqueluche

Depuis l'introduction des programmes de vaccination dans les années 1940 le nombre de cas de coqueluche a diminué significativement. Cependant la coqueluche reste parmi les 10 maladies infectieuses causant le plus haut taux de morbidité et mortalité de par le monde. Plus inquiétant encore, dans les pays à haute couverture vaccinale la coqueluche reste endémique et certains de ces pays ont vu le nombre de cas augmenter depuis le début des années 1980 ce qui qualifie la coqueluche de maladie infectieuse à prophylaxie vaccinale la plus répandue dans ces pays (Mooi et al., 2009).

Les premières références littéraires mentionnant la coqueluche datent du XVI^{ème} siècle avec notamment une description d'une épidémie survenue à Paris en 1578 (Lapin, 1943). Il est possible que le contact entre *Bordetella pertussis* et l'homme soit assez ancien et que le caractère monomorphe de *Bordetella pertussis* soit dû à la dissémination d'un clone en particulier favorisée par l'amplification récente de la migration humaine. Alternativement le caractère aigu de l'infection a pu apparaître récemment grâce à l'acquisition par *Bordetella pertussis* d'une capacité de provoquer des symptômes plus sévères.

Le genre *Bordetella* appartient à la famille des *Alcaligenaceae* qui inclut également les genres *Alcaligenes* et *Achromobacter* (Brenner, 2005). Le genre *Bordetella* comprend à ce jour 9 espèces : *B. bronchiseptica*, *B. pertussis*, *B. parapertussis*, *B. Holmesii*, *B. petrii*, *B. avium*, *B. hinsii*, *B. trematum*, *B. ansoipii*. Une étude des relations phylogéniques en utilisant un algorithme du 'plus proche voisin' sur les séquences du gène de l'ARN ribosomique 16S montre que *B. bronchiseptica*, *B. pertussis*, *B. parapertussis* et *B. Holmesii* forment une lignée distincte et regroupée clairement séparée des autres membres du genre.

1.2 L'évolution des Bordetelles

B. pertussis, *B. bronchiseptica* et *B. parapertussis* sont considérés comme génétiquement extrêmement proches les uns des autres sur base d'études d'hybridation comparatives des génomes, de l'identification et de l'étude du polymorphisme des différents IS, d'études de polymorphisme de restriction multi-locus (MLEE) et de caractéristiques métaboliques (Musser, 1986, Stibitz, 1989, Zee, 1997).

Les relations phylogéniques entre les Bordetelles "classiques", *Bordetella bronchiseptica*, *Bordetella parapertussis* et *Bordetella pertussis*, ont été plus particulièrement étudiées comparativement aux autres espèces en utilisant des techniques de MLEE (Musser et

al 1986, 1987). Ces études ont révélées que la diversité entre ces souches était en fait insuffisante pour les classer comme souches différentes et que cette classification n'avait de justification que médicale et historique. D'un point de vue phylogénétique, *Bordetella pertussis* et *Bordetella parapertussis* ne sont que des clones de *Bordetella bronchiseptica* spécialisés à l'humain. Cette diversité génétique limitée a été expliquée soit par une origine récente soit par un goulot d'étranglement génétique.

L'étude de la distribution des différentes IS dans la population a montré une bonne corrélation avec les études phylogéniques basées sur les MLEE et en font des cibles PCR de choix pour les diagnostics.

Récemment les relations phylogénétiques des Bordetelles classiques ont été encore précisées par des approches de typage par séquençage multilocus (MLST) et d'hybridation comparative (CGH) (Diavatopoulos, 2005). L'arbre couvrant minimum obtenu à partir de ces approches permet de distinguer 4 complexes, les complexes I et IV contient 2 lignées différentes de *Bordetella bronchiseptica*, le complexe II contient les *Bordetella pertussis* et le complexe III contient les *Bordetella parapertussis*_{HU} (*Bordetella parapertussis* infectant les humains). Un complexe est formé de souches plus proches entre elles que des souches en dehors du complexe. Le moment de divergence entre *Bordetella pertussis* et les souches du complexe IV de *Bordetella bronchiseptica* est estimé entre 0,32 et 2,53 millions d'années.

Les séquençages des génomes de ces souches ont permis d'établir le schéma évolutif du genre Bordetelle qui prend source au niveau d'un ancêtre '*B. bronchiseptica*-like' et qui suggère l'apparition d'une lignée *B. pertussis* et *B. parapertussis* au départ de cet ancêtre commun. L'évolution s'est probablement faite ensuite principalement par remodelage, réarrangement et perte des certaines parties du génome pour atteindre un mode de vie sans étape environnementale et, dans le cas de *B. pertussis*, une pathogénie stricte de système respiratoire de l'homme. Cette évolution est principalement due à l'action des 'Insertion Sequences' (IS) qui permettent une plasticité du génome entraînant des réarrangements de grandes régions du génome, l'inactivation par insertion de ces IS dans les ORF de gènes non-essentiels au mode de vie de *B. pertussis* et enfin l'excision de régions entières bordées par des IS par mécanisme de 'looping out'. Cet état de fait est admirablement traduit par deux observations essentielles. La première concerne la comparaison des gènes prédits dans les 3 grandes souches et des gènes en communs (Figure 1.1) (Parkhill et al 2003, Nat Genetics).

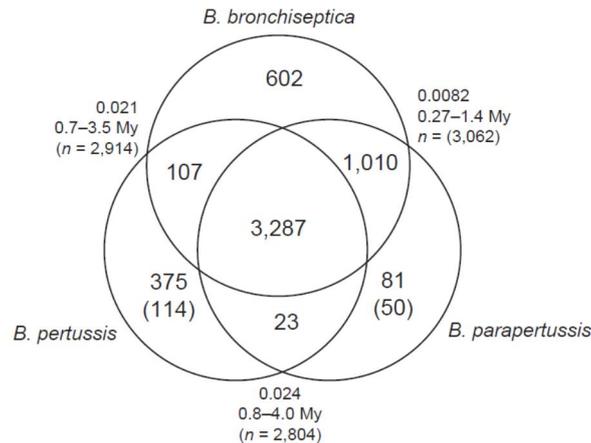


Figure 1.1 : Diagramme de Venn présentant les nombres de gènes uniques à *B. pertussis* Tohama I, *B. bronchiseptica* RB50 et *B. parapertussis* 12822 et commun à 2 des 3 ou aux 3 souches. Les chiffres entre parenthèses correspondent aux gènes uniques une fois les IS exclus. Les valeurs en dehors des cercles correspondent au taux moyen de substitutions synonymes ainsi que l'âge estimé de divergence entre les souches (d'après Parkhill *et al.* 2003).

B. bronchiseptica (RB50) présente 5007 gènes, *B. parapertussis* (12822) 4404 et *B. pertussis* (TohamaI) 3816. Parmi ces gènes 3287 sont communs aux 3 souches, un nombre important, 1010, sont communs à *B. bronchiseptica* et *B. parapertussis* mais pas présents chez *B. pertussis*. Par ailleurs un faible nombre, 23, sont communs entre *Bordetella pertussis* et *B. parapertussis* mais pas chez *B. bronchiseptica* et 107 sont communs à *B. pertussis* et *B. bronchiseptica* mais pas chez *B. parapertussis*. Illustrant l'évolution indépendante de ces souches à partir d'un ancêtre commun *B. bronchiseptica*-like par réduction du génome. Il n'y a pas ou pratiquement pas eu d'acquisition de gènes de *B. pertussis* ou *B. parapertussis* provenant de *B. bronchiseptica*. La deuxième observation est obtenue par comparaison linéaire des génomes (figure 1.2) en BLASTN qui montre nettement une dynamique de réarrangement importante chez *B. pertussis* par rapport à *B. bronchiseptica* qui est beaucoup moins marquée chez *B. parapertussis*.

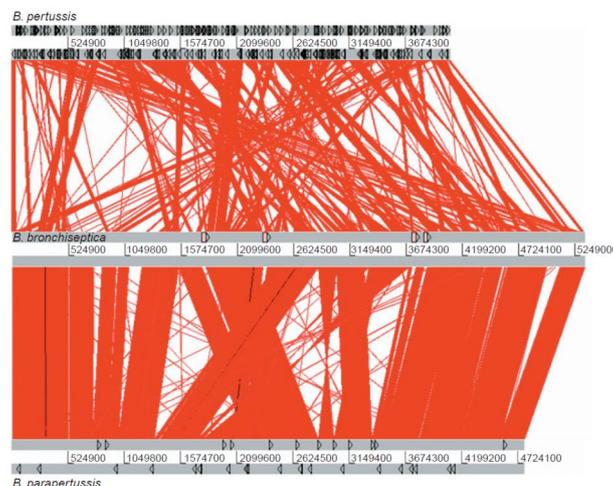


Figure 1.2 : Comparaison des compositions génomiques des souches *B. pertussis* TohamaI, *B. bronchiseptica* RB50, *B. parapertussis* 12822. Les IS sont représentées par des triangles noirs. Les régions génomiques présentant des similarités par analyse BLASTN sont reliées par les lignes rouges (d'après Parkhill *et al.* 2003).

Cet ensemble d'observations et leurs ampleurs est strictement proportionnel au nombre d'IS dans les différents génomes 361 chez *B. pertussis*, 112 chez *B. parapertussis* et 0 chez *B. bronchiseptica*. L'origine de ces IS, le moment de leur acquisition par la bactérie et leur mode de dissémination dans les différentes souches de Bordetelles est controversé. Il est à noter que la souche 18323 est une souche à part, elle contient des régions génomiques présentes chez *B. bronchiseptica* mais absentes chez d'autre *B. pertussis*. Par ailleurs il lui manque 9 gènes qui sont spécifiques des autres *B. pertussis* par rapport à *B. bronchiseptica* et *B. parapertussis* suggérant une évolution particulière de cette souche. On peut la considérer soit en tant qu'intermédiaire entre *B. bronchiseptica* et *B. pertussis* soit résultant d'un processus évolutif séparé des autres.

Il a été émis l'hypothèse que les Bordetelles ont subi deux pressions de sélection lors de l'évolution de *B. bronchiseptica* à *Bordetella pertussis* et *Bordetella parapertussis* (Bjornstad et Harvill, 2005) : (i) une sélection pour changer d'un mode d'infection chronique à un mode d'infection aiguë, (ii) une sélection pour échapper à l'immunité de l'hôte.

L'évolution d'un pathogène fortement contagieux mais peu infectieux vers un pathogène faiblement contagieux mais fortement infectieux est un thème récurrent dans le monde des pathogènes. Les pathogènes dont la population d'hôte est peu dense tendent à avoir une pathologie chronique qui permet une colonisation longue de l'hôte, augmentant ainsi les chances de rencontrer un nouvel hôte susceptible. Ces pathogènes sont généralement également capables de survivre dans l'environnement facilitant ainsi l'acquisition par un nouvel hôte mobile. Une population d'hôtes plus dense (urbanisation) augmente la fréquence de rencontre d'un hôte susceptible et favorise l'apparition d'infections aiguës avec des symptômes particuliers (éternuements, toux, ...) qui favorisent la dissémination d'hôte à hôte sans étape dans l'environnement. L'urbanisation au 16^{ème} siècle correspond à l'apparition des premiers rapports d'épidémies de coqueluche.

Les effets de ces évolutions au niveau du génome se traduisent par la réduction de la taille du génome par élimination des gènes non essentiels à un mode de vie restreint à l'homme. En effet les gènes présents chez *Bordetella bronchiseptica* et pas chez *Bordetella pertussis* ou *Bordetella parapertussis* sont principalement les protéines de surface et leurs systèmes qui probablement permettaient l'interaction avec l'environnement. En ce qui concerne le passage d'une infection chronique à aiguë les différences se situent au niveau de l'interaction avec le système immunitaire de l'hôte. Par exemple la Toxine (PT) est détectée chez *Bordetella pertussis* mais pas chez *Bordetella bronchiseptica* et *Bordetella parapertussis*. Or, chez la souris, PT inhibe le recrutement des neutrophiles

dans le tractus respiratoire pour retarder l'élimination médiée par les anticorps (=> infection aiguë). La différence d'expression entre les Bordetelles est due à un pseudogène au niveau de *ptxB* chez *Bordetella parapertussis* et un promoteur présentant moins d'affinité pour l'ARN-polymérase et pour BvgA chez *Bordetella bronchiseptica* que chez *Bordetella pertussis*. Un autre critère se situe au niveau de la stimulation des Toll-like receptor 4 (TLR-4) par les lipides A qui est critique pour la limitation du nombre de bactérie dans le modèle d'infection murin. *Bordetella pertussis* et *Bordetella parapertussis* ont développé des LPS présentant une stimulation des TLR-4 *in vitro* 10x à 100x moindre que chez BB (Mann et al., 2004 ; Mann et al., 2005).

1.3 Effet de la vaccination sur l'évolution de *Bordetella pertussis*

Les collections de souches disponibles pour *B. pertussis* datant de l'aire pré- ou post-vaccinale permettent une étude approfondie de l'évolution à court terme d'un pathogène face à une pression vaccinale. Ces études ont permis de montrer une réduction de la variabilité génétique au niveau de la population à la suite de l'introduction des programmes de vaccination probablement due à une réduction de la taille de la population suivie d'une expansion de quelques clones particuliers (van Loo et al., 1999; van Loo et Mooi, 2002; Elomaa et al., 2005; Weber et al. 2001). D'une manière générale, une diversité génétique très faible a été constatée dans la population de *B. pertussis* en comparaison avec d'autres bactéries pathogènes respiratoires (van Loo et al., 2002; Packard et al., 2004). Les souches circulantes récentes ne présentent des polymorphismes codants que dans 5 des 15 gènes étudiés précisément. Il s'agit des gènes de la pertactine (*prn*), de la sous-unité 1 de la toxine pertussique (*ptxA*), du facteur A de la colonisation trachéale (*tcfA*) et des fimbriae 2 et 3 (*fim2* et *fim3*).

Dans le gène de la pertactine les variations surviennent dans les régions appelées R1 et R2. Ces régions sont assez éloignées l'une de l'autre dans la séquence primaire mais probablement proches dans la protéine repliées (Hijnen et al. (2004)) formant ainsi des épitopes qui ont été montrés comme potentiellement protecteurs (Charles et al., 1991; King et al., 2001). La toxine pertussique est aussi le lieu de variations qui touchent principalement PtxA, la sous unité portant l'activité toxique (van Loo et al., 2002). Des études ont montré que les variations observées survenaient dans un épitope reconnu par les cellules T (De Magistris et al., 1992) et au niveau d'un site de fixation d'un anticorps monoclonal neutralisant (Bartoloni et al., 1988). Cependant l'impact de ces modifications sur l'habilité des anticorps à reconnaître et neutraliser la toxine reste controversé (Hausman et Burns, 2000). Les fimbriae ont fait l'objet de nombreuses études du fait de leur implication dans l'immunité protectrice (Cherry et al., 1998;

Storsaeter et al., 1998). Chez *B. pertussis* seuls *fim2* et *fim3* sont exprimés et confèrent le sérotype 2 et/ou 3 aux souches circulantes (Willems et al., 1992; Lochter et al., 1992; Parkhill et al., 2003). Cette expression est fonction de la longueur d'une répétition homo-polymérique de C dans la séquence du promoteur de ces gènes (Willems et al., 1990). Un polymorphisme sur *fim2* et 3 sur *fim3* ont été dénombrés résultant dans la circulation de 2 allèles de *fim2* et 4 allèles de *fim3* différentes (van Loo, et al. 2002; Tsang et al., 2004).

La circulation des différents allèles de *prn*, *ptxA* et des *fim* et l'impact des allèles utilisés dans la composition des différents vaccins acellulaires de par le monde sur les souches circulantes ont été étudiés. Un changement dans la fréquence d'apparition d'allèles en fonction de la composition vaccinale a pu être mis en évidence (Mooi et al., 1998; Mooi et al., 2001).

1.4 La virulence et sa régulation chez *Bordetella pertussis*

1.4.1 Système BvgAS

Très tôt dans l'étude moléculaire du pathogène *B. pertussis* il a été montré que cette bactérie présentait une capacité de modulation de sa virulence en fonction des conditions de culture (Lacey, 1960). Une phase virulente présentant une expression maximale des antigènes de virulence (appelée plus tard phase bvg+) a été ainsi définie et a été opposée à une phase avirulente (phase bvg-). Cette modulation phénotypique correspond probablement à une réponse adaptative de la bactérie en fonction de signaux environnementaux (voir ci-dessous).

Chez les Bordetelles un très grand nombre de gènes est sous le contrôle du locus *bvgAS*. Cet opéron code pour un système à deux composants dont BvgS est le senseur ancré dans la membrane interne de la bactérie et BvgA le régulateur cytoplasmique, (Arico et al., 1989 ; Stibitz et Yang, 1991). Le senseur, BvgS, présente un domaine périplasmique qui est probablement impliqué dans la perception d'un signal environnemental déterminant pour la régulation de l'état de virulence de la bactérie. Cette perception de signal est transmise à l'intérieur de la cellule pour activer le domaine cytoplasmique qui est composé (en partant de la membrane) d'un 'linker' reliant la région transmembranaire au 1^{er} domaine, le 'domaine kinase', lui-même lié au domaine receveur à son tour relié le domaine Hpt (Histidine Phospho-Transfert domaine). Lors d'une perception de signal, le domaine périplasmique est activé, il va provoquer l'autophosphorylation du domaine kinase sur un résidu kinase conservé (voir figure 1.3). Une étude récente de la structure de BvgS montre l'importance de la conformation entrelacée des domaines VFT (Venus FlyTrap) de la portion

periplasmique qui semble exercer une contrainte sur la motié cyroplasmique afin de la maintenir une configuration 'kinase-on' (Dupré *et al.*, 2015).

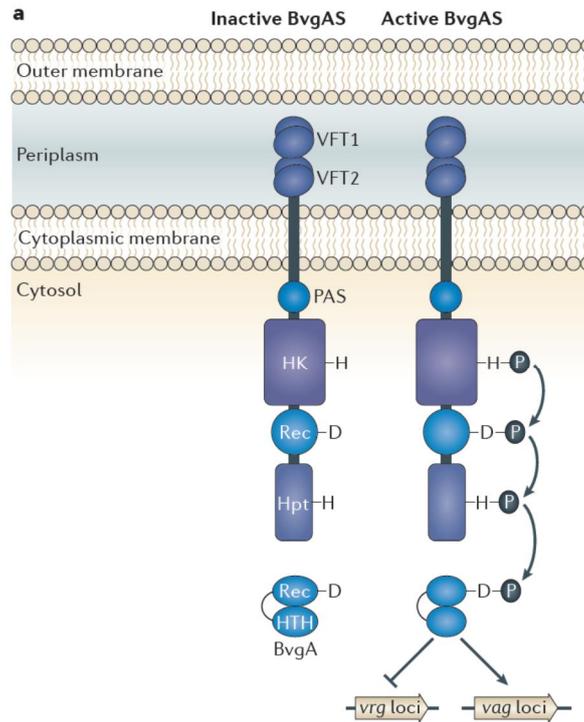


Figure 1.3 : Représentation schématique des domaines de BvgS. La cascade de phosphorylation entre les différents domaines lors de l'état activé du récepteur est représenté des P cerclés. BvgA est l'accepteur final de la cascade de phosphorylation. (d'après Melvin *et al.* 2014).

Le phosphate va être ensuite transféré sur un résidu aspartate du domaine receveur puis à une kinase du domaine Hpt. L'accepteur final de cette cascade de phosphorylation est un résidu aspartate conservé dans le domaine receveur de BvgA (Uhl et Miller, 1994 ; Uhl et Miller, 1996a ; Uhl et Miller, 1996b).

Il est important de noter que, contrairement aux autres systèmes à deux composants bactériens, le système BvgAS semble être actif par défaut (dans les conditions normales de culture ou *in vivo*). A ce jour le signal environnemental provoquant l'activation du système n'a pas été identifié et seul des modulateurs connus pour inactiver la cascade de phosphorylation ont été mis en évidence. Une présence en concentration importante de MgSO₄ ou d'acide nicotinique, ou encore une température de culture réduite, provoque cette inactivation du système.

L'effet exact de la phosphorylation de BvgA (BvgA-P) sur sa structure et sur son mode d'activation n'est pas encore entièrement clair. Le scénario le plus souvent évoqué est que la phosphorylation favorise la dimérisation de BvgA qui va alors entrer en interaction avec le promoteur des gènes activés lors de la virulence (*vag* : 'virulence activated genes') et activer la transcription de ces gènes.

1.4.2 Etude des motifs des promoteurs de gènes

Le mode de fixation et le processus d'activation par BvgA n'ont été étudiés que sur quelques gènes. Les promoteurs de deux gènes ont été étudiés en particulier, celui de *fhaB* et celui de *ptx*. Dans les années 1980 des études en 'footprinting' de BvgA surexprimé dans *E. coli* ont permis d'établir l'importance d'une petite région centrée en position -88,5 pb par rapport au départ de transcription. Ce site contient une heptade inversée quasi-parfaite (Roy et Falkow, 1991). Il fut établi ensuite que ce site correspond au site primaire de fixation de BvgA-P et qu'il existe des sites secondaires entre cette position et la région -35 (Barcher et al., 1997). La fixation sur ces sites secondaires se fait de manière coopérative par oligomérisation avec le BvgA fixé sur le site primaire plus que par reconnaissance de motif de séquence (Boucher et al, 2001a). L'ARN polymérase joue également son rôle dans la formation et la stabilisation du complexe tertiaire : séquence promotrice/BvgA-P/RNA-pol (Jones et al, 2005).

Ces mêmes études ainsi que d'autres études sur *ptx* et sur d'autres gènes de virulence ont établi que lors d'un passage d'une phase *bvg-* à *bvg+* (passage d'une faible température 20-22°C à 37°C, par exemple) certains promoteurs, comme celui de *fhaB*, était activé rapidement alors que d'autres, tel que celui de *ptx*, l'étaient plus tardivement (Scarlatto et al, 1991). Ces différents modes de réponse ont été désignés 'early' et 'late' promoteur et une série de gènes *bvg*-dépendants ont été classifiés comme 'early' ou 'late genes'. L'activation d'un promoteur 'late' nécessite plus de BvgA-P dans la cellule. Ce mode de régulation temporel correspondrait à un besoin rapide, lors de l'infection, de facteurs d'adhésion tel que FHA et d'un besoin plus tardif d'autres facteurs de virulence telle que l'action des toxines pour détourner la réponse immunitaire de l'hôte et causer des lésions dans l'épithélium pulmonaire.

En dehors des nombreux *vag* on dénombre également un certain nombre de *vrg* ('virulence repressed genes') dont l'expression maximale est caractéristique de la phase avirulente de la bactérie (phase *bvg-*). La régulation de ces *vrg* a longtemps été mise sous le contrôle de d'un facteur appelé BvgR (dont le gène a un profil d'expression de type *vag*). BvgR agirait comme répresseur des gènes *vrg* lors de la phase *bvg+* et leur expression serait dé-réprimée suite à l'absence de BvgR lors de la phase *bvg-*. Cependant l'étude de la séquence de BvgR n'a montré aucune caractéristique connue de protéine-liant l'ADN mais plutôt la présence d'un domaine phosphodiésterase EAL (D'Argenio and Miller, 2004). Ce type de domaine est impliqué, en conjonction avec un domaine de type GGDEF, pour la régulation

phénotypique de nombreuses bactéries *via* un second messager de type cyclic-di-GMP (c-di-GMP). Il est donc probable que BvgR exerce sa régulation sur les *vrg* *via* un autre facteur sensible à la concentration en c-di-GMP et non pas en fixant directement les régions promotrices des *vrg*.

1.4.3 Profils d'expression

L'ensemble des profils de modulation des gènes de virulence chez *Bordetella pertussis* a été résumé sous forme d'un diagramme par Peggy Cotter (Cotter et al., 2003). On distingue 4 profils de modulation de gènes en fonction de la concentration en BvgA phosphorylé (BvgA-P) dans la cellule. Ces profils sont dénommés : 'vrg', les 'vig', les 'early vag' et les 'late vag' (voir figure 1.4)

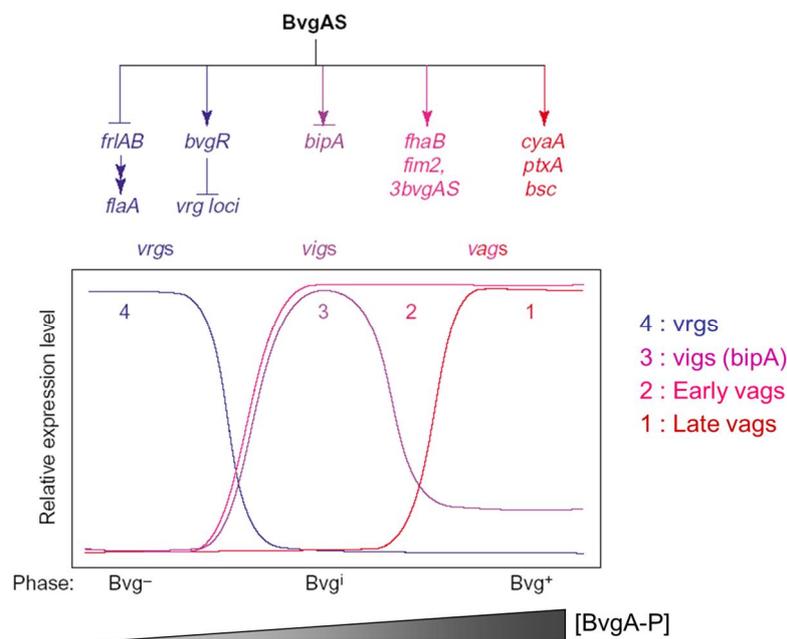


Figure 1.4 : Profils d'expression des différentes classes de gènes en fonction de la phase de virulence bvg-, bvgi, bvg+, elles même fonction de la concentration en BvgA-P (d'après Cotter et al., 2003)

1.5 Les principaux gènes de virulence : adhésines et toxines

1.5.1 Les adhésines

Les Bordetelles ont développé un réseau important de protéines d'adhésion associées à leurs systèmes de production à la surface. La fonction principale de ces adhésines réside dans l'adhérence efficace et ciblée à différentes cellules assurant ainsi la colonisation de l'hôte infecté. Cependant d'autres implications de ces adhésines dans la virulence de *Bordetella pertussis* ont été documentées, elles participent notamment à l'altération de la réponse immunitaire de l'hôte et elles représentent de par leur exposition à la surface et leur abondance, des cibles immunitaires de première

importance. A ce titre, certaines de ces adhésines rentrent dans la composition des antigènes protecteurs que l'on retrouve dans les vaccins acellulaires.

Toutes les adhésines connues de *Bordetella pertussis* sont sous le contrôle du système à deux composants BvgAS. C'est en étudiant la régulation de transcription de FHA, l'hémagglutinine filamenteuse, et en la comparant à l'expression d'autres vags qu'il a été établi l'existence de vags 'early' et 'late' (voir paragraphe 1.4.2) (Scarlato et al 1991 ; Stibitz, 1998a). Le gène *fhaB* a été montré comme étant transcrit très rapidement après infection chez la souris. Quelques heures plus tard c'est au tour de la pertactine d'être exprimée avec la toxine pertussique suivi encore plus tard par l'adénylate cyclase. Cette régulation temporelle précise permet l'adhérence initiale ciblée et la succession optimale d'interactions avec les cellules de l'hôte y compris au niveau des cellules de la réponse immunitaire (Veal-Carr et Stibitz 2005). L'altération de cette expression séquentielle par interversion des promoteurs 'early' et 'late' a montré une colonisation affectée chez la souris (Kinnear et al 2001).

1.5.1.1 Fimbriae

Les fimbriae sont des structures péritriches à la surface externe des bactéries. Elles sont formées d'un grand nombre de sous unités fimbriales associées entre elles sous forme de filaments et à la pointe duquel on trouve la protéine FimD (Geuijen et al 1997). Chez *Bordetella pertussis* on trouve 2 sous-unités de fimbriae majeures *fim2* et *fim3*. Deux autres gènes sont connus : *fimA* qui est tronqué chez *Bordetella pertussis* uniquement (Boschwitz et al 1997b ; Willems et al 1992) et *fimX* qui est transcrit et régulé par BvgAS mais apparemment non exprimé (Pedroni 1988, Riboli 1991).

L'expression de *fim2* et *fim3* est sous le contrôle d'une boîte de 15 bases C dans leur promoteur qui peuvent subir des transitions de phase par petites insertions ou délétions dans ce stretch de C résultant en une faible ou une forte transcription du gène (Willems 1996). Les fimbriae sont impliqués dans la colonisation et la persistance lors de l'infection chez la souris (Mooi, 1992). Les fimbriae semblent interagir avec FHA lors de la colonisation de la trachée chez la souris (Kimura, 1990) et sont impliqués dans l'adhérence de *Bordetella pertussis* au niveau des cellules de l'épithélium laryngale (Van den Berg, et al 1999). Il semblerait que les fimbriae peuvent avoir un effet anti inflammatoire (Vandebriel et al 2003).

1.5.1.2 FHA, l'hémagglutinine filamenteuse

L'hémagglutinine filamenteuse est produite à la surface des bactéries sous forme d'un agrégat de matériel filamenteux. Ce matériel correspond à l'agrégation d'une partie Nt du produit du gène *fhaB*. La sécrétion et l'assemblage sont assurés par un système de

sécrétion à deux partenaires : TPS (two partners secretion) (Clantin et al, 2004 ; Jacob Dubuisson et al 2001) qui représente le paradigme de ce type de système. Le précurseur de FHA est maturé à la surface de la cellule par SphB1 une protéase spécifique de la famille des subtilisine (Coutte et al, 2001).

La FHA est une adhésine majeure, elle a été montrée comme capable de fixer toute une série de cellules différentes *in vitro* (Urisu et al 1986 ; Arico et al 1993 ; Hannah et al 1994 ; Van den Berg et al 1999a 1999b ; Saukkonen et al 1991 ; Hellwig et al 1999 ; Hazenbos et al 1993) et des expériences chez l'animal montrent son importance dans la colonisation du tractus respiratoire (Alonso et al 2001 ; Kimura et al 1990) et dans la persistance de la bactérie (Khelef et al 1994, Melvin et al., 2015).

1.5.1.3 Pertactine

La pertactine est un autotransporteur de type 1 (Jacob-Dubuisson et al., 2004). Sa maturation se fait par clivage de l'extrémité Ct du polypeptide qui forme un domaine de translocation dans la membrane au travers duquel le 'passenger domain' est externalisé. Ce domaine est formé d'hélices β similaires à celles de la FHA. Le rôle exact et le récepteur de fixation de la pertactine *in vivo* sont encore inconnus mais la pertactine interviendrait dans la résistance à la clearance par les neutrophiles. La pertactine présente des polymorphismes dans des épitopes conformationnels (Hijnen et al, 2004). Plusieurs variants appelés prn1, 2 et 3 circulent. Un shift du variant 1 (que l'on trouve dans les souches vaccinales) vers les variants 2 et 3 a été observé dans les souches circulantes ces dernières années, ce qui peut expliquer en partie l'évasion immunitaire observée (He et al, 2003 ; Hijnen et al, 2004).

1.5.2 Les toxines

La coqueluche a d'abord été considérée comme une maladie médiée uniquement par les toxines de la bactérie. On sait maintenant que la pathogénèse est complexe et fait intervenir les toxines, les adhésines et d'autres composants bactériens.

1.5.2.1 La toxine pertussique

La toxine pertussique est une exotoxine du type AB₅, le domaine A possède la fonction catalytique et le pentamère B assure le transport et la fixation aux récepteurs (Stein et al, 1994a). Elle est encodée par les gènes *ptxA-E* et est sécrétée par un système de sécrétion de type V lui-même encodé par les gènes *ptlA-I* (Kotob et al, 1995). La toxine fixe toutes les glycoprotéines présentant de l'acide sialique elle a donc une multitude de récepteurs possibles et peut donc agir sur différents types cellulaires (Stein et al, 1994b ; Loch et al., 2011 (FEBS)). Après fixation sur le récepteur le complexe est

endocyté et subit un transport rétrograde jusqu'au Golgi puis au réticulum endoplasmique (el Baya et al., 1999). La sous-unité A sort du réticulum endoplasmique probablement *via* le pathway de dégradation des protéines malformées (Worthington et al., 2007). La sous-unité A transfère alors l'ADP ribose du NAD⁺ vers des protéines G hétérotrimériques y compris vers des protéines G inhibitrices. L'inhibition de l'adénylate cyclase ne se fait alors plus et provoque l'augmentation de l'AMP cyclique (AMPc) intracellulaire ce qui a pour conséquence une dérégulation de la réponse immunitaire (Locht et al., 2011 (FEBS) ; Graf et al., 1992).

Le rôle exact de la toxine pertussique lors de l'infection n'est pas encore précisément déterminé du fait de la variété des effets observés en culture cellulaire et en modèle animal. *In vitro* cette toxine inhibe la migration des cellules qui expriment des récepteurs couplés aux protéines G, telles que les neutrophiles, les monocytes et les lymphocytes (Spangrude et al., 1985). Dans des modèles murins, la production de la toxine pertussique est en corrélation avec la diminution de cytokines et de chimiokines pro-inflammatoires ce qui diminue le recrutement des neutrophiles vers les poumons et augmente les charges bactériennes au début de l'infection (Andreasen et al., 2008 ; Kirimanjeswara et al., 2005). Des expériences chez la souris montrent que les cellules principalement visées sont les macrophages alvéolaires (Carbonetti et al., 2007).

1.5.2.2 L'adénylate cyclase

L'adénylate cyclase est codée par le gène *cyaA*. C'est une toxine de la famille des RTX (Repeats in Toxine) sécrétée par un système de sécrétion de type I, lui-même codé par les gènes *cyaBDE*. On distingue dans la structure de la toxine 2 domaines fonctionnels, le domaine Ct qui contient les répétitions RTX et qui permettent la reconnaissance et la fixation des cellules cibles pour former des pores à cations dans leurs membranes (El-Azami-El-Idrissi et al., 2003 ; Sakamoto et al., 1992) et le domaine Nt qui est une adénylate-cyclase calmoduline dépendante qui va convertir l'ATP en AMPc après translocation dans la cellule cible (Ladant et al., 1989 ; Glaser et al., 1989). Les effets combinés de l'intoxication à l'AMPc et de la formation de pores inhibent la phagocytose par le complément (Kamanova et al., 2008), induit la production de cytokines anti-inflammatoires, supprime la production de cytokines pro-inflammatoires et enfin inhibe le recrutement des cellules immunitaires (Confer et al., 1982).

Tout comme la toxine pertussique, l'adénylate cyclase est capable d'intoxiquer beaucoup de types cellulaires différents. Elle reconnaît et fixe le récepteur C3 du complément (aussi appelé intégrine $\alpha_M\beta_2$) que l'on retrouve sur les neutrophiles, les macrophages et les cellules dendritiques (Guermonprez et al., 2001). Des expériences

chez la souris suggèrent que l'adénylate cyclase joue un rôle majeur dans la résistance de la bactérie à la clearance par les neutrophiles (Harvill et al. 1999 ; Henderson et al. 2012).

1.6 Introduction aux travaux de recherches

La persistance et la réémergence des cas de coqueluche malgré les efforts de vaccination est difficilement compréhensible du fait du caractère particulièrement clonal et du manque de diversité génétique de ce pathogène (Caro et al., 2006a ; Diavatopoulos et al., 2005 ; Parkhill et al., 2003 ; van Loo et al., 2002). Les différences observées entre les isolats cliniques sont principalement attribuées à une expression différentielle des gènes de protéines de surface, des mutations dans les protéines sécrétées et une différence de taille des génomes *via* l'action de séquences d'insertion (Brinig et al., 2006 ; Caro et al., 2006b ; Heikkinen et al., 2007). Contrairement à d'autres pathogènes comme les salmonelles ou les bactéries du genre *Vibrio* les différences de virulence ne peuvent être attribuées à des îlots de pathogénie, des plasmides ou des insertions d'origine phagique. Il semblerait par contre que le régulon de virulence, qui est large et complexe chez *B. pertussis* (Hot et al., 2003, Cummings et al. 2006), soit différemment exprimé dans les différentes sous-espèces expliquant la diversité des niches et des modes de vie (Cotter et al., 2000, Mattoo et al., 2000). Il a été montré qu'un changement dans le mode d'expression des gènes de virulence peut réduire significativement la colonisation chez la souris (Akerley et al. 1995 ; Kinnear et al., 2001) mettant en évidence l'importance de la régulation des réseaux de gènes sur le caractère virulent de la bactérie.

Notre recherche est axée sur cette constatation et tente d'établir le lien entre la génomique au sens large de la bactérie (niveau de transcription et régulation, composition et organisation génomique, cartographie du transcriptome et découverte de nouveaux transcrits, petits ARN régulateurs et leurs fonctions) et sa virulence. Les projets que nous avons mis en place visent à éclaircir le mode de régulation de la virulence chez la bactérie et comment cette régulation pourrait intervenir dans la résurgence actuelle du pathogène. Nous présentons ici plusieurs exemples de résultats :

- (i) Une étude transcriptomique à haut débit par le développement et l'utilisation d'une puce à ADN spécifique de *Bordetella pertussis* (1^{ère} en son temps) qui a permis de mettre en évidence de nouveaux *vag* et d'affiner notre compréhension de la régulation par le système à 2 composants BvgAS
- (ii) Une étude en CGH-array, avec une puce à ADN étendue au génome entier, qui a permis de caractériser finement les groupes de PFGE en termes de composition génomique et ainsi d'étudier l'évolution des souches circulantes par détermination de leur composition génomique.

- (iii) Une première étude d'identification et de caractérisation de petits ARN régulateurs chez les Bordetelles.
- (iv) Une étude non encore publiée du transcriptome primaire de *Bordetella pertussis* et des caractéristiques importantes de ce transcriptome avec notamment l'établissement d'une liste de nouveaux transcrits.
- (v) Deux études de l'importance du facteur Hfq dans la régulation de la virulence chez *Bordetella pertussis* donnant un aperçu de l'implication des petits ARNs non codants dans la régulation des mécanismes de virulence.
- (vi) Enfin, la caractérisation en cours d'un petit ARN non codant, BprC, qui régule probablement le niveau de traduction de *bvgA* soit en stabilisant l'ARNm soit en favorisant l'entrée des ribosomes sur le RBS.

TRAVAUX DE RECHERCHE

CHAPITRE 2 : ETUDE DU REGULON DE BVGAS

Annexe 1 : **Hot, D.**, Antoine, R., Renauld-Mongenie, G., Caro, V., Hennuy, B., Levillain, E., Huot, L., Wittmann, G., Poncet, D., Jacob-Dubuisson, F., Guyard, C., Rimlinger, F., Aujame, L., Godfroid, E., Guiso, N., Quentin-Millet, M. J., Lemoine, Y., and Locht, C. (2003). Differential modulation of *Bordetella pertussis* virulence genes as evidenced by DNA microarray analysis. *Mol Genet Genomics* 269(4), 475-86.

Au début des années 2000 la connaissance du régulon du système à deux composant BvgAS de *Bordetella pertussis* n'était que partielle. BvgS et BvgA avaient été identifiés comme le senseur et l'accepteur respectivement du système à 2 composants et l'importance de ce système dans la régulation de la virulence avait été clairement établie par l'étude de la régulation de plusieurs toxines, adhésines, protéases et autotransporteurs. Cependant l'étendue de l'influence de BvgA en termes de nombre de gènes sous sa dépendance et de la diversité de fonctions de ces gènes était insoupçonnée. Deux développements majeurs à cette époque vont permettre une étude beaucoup plus complète et précise du régulon BvgAS. Tout d'abord la mise disposition du public par le Sanger Institute du séquençage du génome de *Bordetella pertussis* Tohamal sous forme d'une séquence non encore annotée mais en un seul contig de taille entière. Cette séquence va permettre de rechercher la présence d'ORF de gènes ou de familles de gènes spécifiques afin de les étudier. L'autre développement majeur réside dans l'émergence à cette époque d'une nouvelle application permettant l'étude en parallèle des niveaux de transcription d'un très grand nombre de gènes. Cette nouvelle application est nommée Puces à ADN (microarray en anglais) et repose sur le dépôt précis, sur une lame de verre, de quelques pico-litres de solution d'ADN dont la séquence a été bioinformatiquement sélectionnée pour être spécifique d'un gène donné. Cet ADN est constitué soit d'oligonucléotides soit de fragments PCR, et la faible quantité des volumes déposés permet de mettre en parallèle des dizaines de milliers de dépôts représentant le génome entier de l'organisme étudié. Une matrice fluorescente est ensuite préparée au départ des ARNm de l'organisme et mise en hybridation sur les fragments d'ADN reporters déposés sur la puce à ADN. Un comparatif des niveaux de fluorescence entre une matrice provenant d'une condition d'étude et une matrice provenant d'une condition de référence permet de calculer les taux de modulation de transcription des gènes entre les deux conditions.

Une première étude, que nous avons initiée en 2001 en collaboration avec l'équipe de Camille Locht (IPL), l'équipe de Nicole Guiso (IP), une équipe de l'ULB et une équipe d'Aventis Pasteur, a été faite sur une sélection de 184 gènes de *Bordetella pertussis* choisis pour leur potentielle implication dans la virulence de la bactérie (voir publication annexe 1). Le choix des régions des ORFs à amplifier s'est fait en travaillant sur l'ébauche de la séquence non-annotée du génome mise à disposition à l'époque sur le site du Sanger Institute. Les gènes sélectionnés comprenaient en plus des facteurs de virulence déjà connus :

- des gènes impliqués dans la capture du fer
- des protéases

- des autotransporteurs
- des facteurs sigmas
- des systèmes à deux composants
- des gènes impliqués dans la glycolyse, le cycle de Krebs, la respiration et la réponse stringente
- des gènes impliqués dans l'utilisation du glutamate et de la proline
- des gènes fixant la pénicilline et des rotamases putatives.

Après une phase de mise au point de la puce à ADN nous avons engagé l'outil développé dans des études des modulations phénotypiques de la souche Tohamal. La bactérie a été cultivée jusqu'en début de phase exponentielle dans le milieu standard Stainer-Scholte (SS) en présence ou absence des agents modulateurs que sont le sulfate de magnésium (MgSO₄, 50 mM) ou l'acide nicotinique (20 mM). Un mutant avirulent délété au niveau de l'opéron *bvgAS* a également été cultivé afin d'étudier son transcriptome. Après culture l'ARN total a été extrait des culots bactériens obtenus et engagés dans des réactions de synthèse de sondes fluorescentes. Les sondes provenant des cultures modulées ou du mutant ont été co-hybridées avec une sonde provenant d'une culture en milieu standard. La différence de niveau de transcription des gènes a ainsi pu être déterminée et exprimé en ratio de modulation entre la condition d'étude et de la condition de référence.

La figure 2.1 représente les ratios obtenus (en valeur logarithmique) des 3 comparaisons pour une série de gènes activés (*vag*) ou réprimés (*vrg*) lors de la phase de virulence.

Cette première exploration du transcriptome de *Bordetella pertussis* a permis tout d'abord de confirmer l'ensemble des *vags* et *vrgs* connu à cette époque et de mesurer le niveau de modulation de ces gènes. De nouveaux gènes ont pu également être classés comme *vags* ou *vrgs*. C'est le cas, pour les *vags*, des autotransporteurs *autoA*, *B*, *D*, *E* de régulateurs ou potentiel régulateurs, *sen5*, *reg11* et *ecf8*, de gènes du système de sécrétion de type III, *bcrJ* et *bscD* et enfin d'un membre de la famille des gènes *bug*. Les nouveaux *vrgs* mis en évidence étaient un autotransporteur *autoC*, un senseur de système à 2 composants *sen2* et un gène de fimbriae *fim3*.

D'une manière plus générale cette étude a montré des niveaux de modulation graduels allant de gènes faiblement modulés jusqu'à très fortement modulés sans pouvoir distinguer de paliers de modulation qui regrouperaient différents gènes. Cette observation traduit le caractère très graduel de l'adaptation du transcriptome de *B. pertussis* à la virulence et le mode de fonctionnement en rhéostat de BvgAS sur son régulon qui sera mis en évidence plus tard par d'autres études sur base de données

issues de microarrays pan-génomiques et dans lequel chaque gène a son profil de réponse à la modulation phénotypique (Cummings *et al.*, 2006).

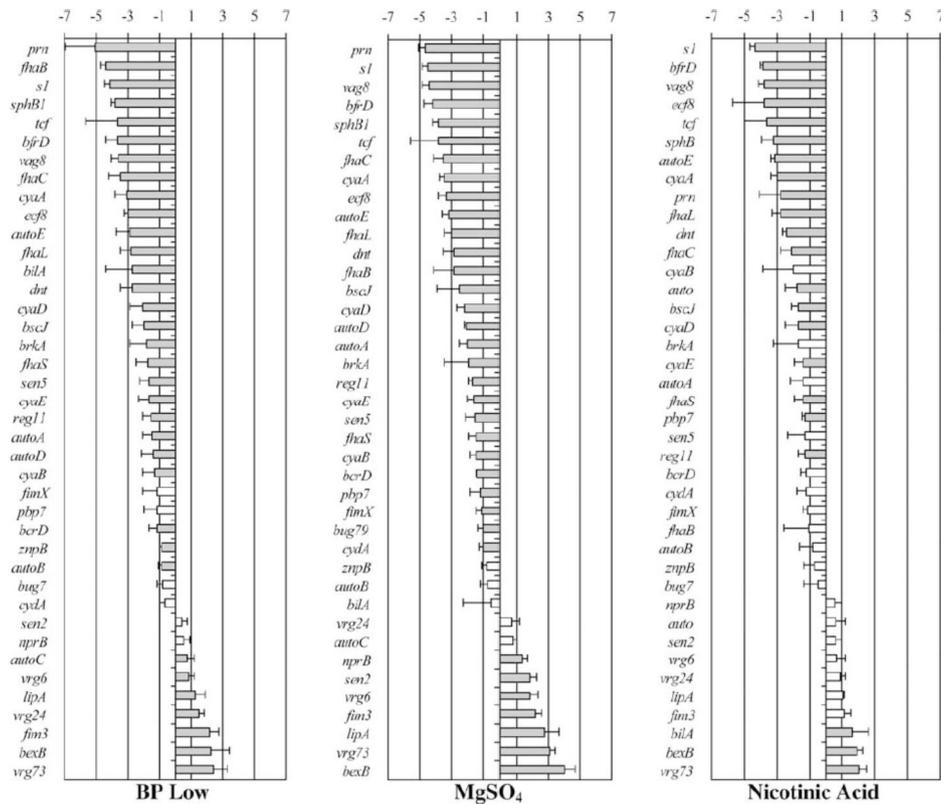


Figure 2.1 : Régulation des vags et vrgs dans 2 conditions de modulation (MgSO₄ et acide nicotinique (« Nicotinic Acid ») et chez le mutant avirulent (BP Low). Les valeurs sont exprimées en log-ratio. (D'après Hot *et al.*, 2003).

Nous avons pu également mettre en évidence par cette étude qu'une poignée de gènes présentait des différences de modulation d'expression entre les 2 conditions de modulation phénotypique (présence de MgSO₄ ou Acide nicotinique) ou chez le mutant BP Low. La grande majorité des gènes qui sont fortement modulés dans l'une des conditions le sont aussi dans les autres conditions mais 4 vrgs (*nprB*, *sen2*, *lipA*, *bexB*) et 2 vags (*fhaB*, *prn*) présentent des différences de niveaux de modulation. Ce résultat laisse entrevoir que des sous-groupes de vags et peut être encore plus fréquemment de vrgs ont des mécanismes de régulation plus complexes que précédemment imaginés. Ces mécanismes font probablement intervenir des régulateurs additionnels ou annexes au régulateur majeur qu'est le système BvgAS. Enfin plusieurs observations ponctuelles sur des gènes importants ont pu être faites. Par exemple les gènes *bscJ* et *bcrD* du système de sécrétion de type III ont été identifiés comme vags forts alors que ces gènes étaient considérés jusqu'alors comme non exprimés (Yuk *et al.*, 1998). Les gènes du système accessoire *cyaBDE* de l'adénylate cyclase (CyaA) sont modulés dans le mutant BP Low ou dans les conditions de modulation phénotypiques alors qu'ils étaient jusqu'alors décrits comme

indépendants de BvgAS (Laoide et al., 1990). Enfin le gène *fim3* a été détecté dans notre étude comme un fort *vrg* alors que la souche utilisée est *fim2⁺ fim3⁻*. L'expression de *fim3* dans la phase bvg- ne semble pas cependant produire de protéine suggérant une régulation post-transcriptionnel pour ce gène. Le pourquoi de cette régulation reste à investiguer.

Quelques années plus tard, lorsque l'annotation du génome fut publiée, nous avons développé une version représentant le génome complet de la bactérie en sélectionnant un reporteur (oligonucléotide de 60-mer) par gène codant (en dehors des transposases des IS). Cet outil est toujours produit de nos jours dans notre laboratoire et sert aux nombreux projets sur *Bordetella pertussis* sur le campus de l'Institut Pasteur de Lille ou en collaboration avec des laboratoires extérieurs.

CHAPITRE 3 : ETUDE DE LA COMPOSITION GENOMIQUE CHEZ *BORDETELLA PERTUSSIS*

Annexe 2 : Caro, V., **Hot, D.**, Guigon, G., Hubans, C., Arrivé, M., Soubigou, M., Renauld-Mongénie, G., Antoine, R., Loch, C., Lemoine, Y and Guiso, N. (2006). Temporal analysis of French *Bordetella pertussis* isolates by comparative whole-genome hybridisation. *Microbes and Infection* 8(8), 2228-35.

Annexe 3 : Lechner, M., Schmitt, K., Bauer, S., **Hot, D.**, Hubans, C., Levillain, E., Loch, C., Lemoine, Y. and Gross, R. (2009). Genomic island excisions in *Bordetella petrii*. *BMC Microbiology*, 9:141.

Le développement d'un microarray représentant le génome entier de *Bordetella pertussis* a permis d'étudier la régulation transcriptomique globale de la bactérie comme décrit et illustré dans le chapitre précédent. Cet outil permet d'investiguer d'autres caractéristiques génomiques notamment avec l'application appelée CGH-array. Cette approche se base sur le marquage puis l'hybridation non pas d'ADNc issu d'une reverse transcription sur les ARNm mais directement du génome lui-même préalablement fragmenté. En hybridant en même temps un génome d'étude (dont on veut déterminer la composition) et un génome de référence (dont la composition est connue) on peut déterminer les parties manquantes et même les parties dupliquées dans le génome d'étude par rapport au génome de référence.

Nous avons appliqué cette approche en 2006 en collaboration avec l'équipe de Nicole Guiso à l'Institut Pasteur afin de caractériser finement les groupes de PFGE établis à l'époque depuis déjà quelques années (Caro et al, 2005 ; Weber et al. 2001). Neuf souches représentatives des 6 groupes différents préalablement déterminés par PFGE (I, II, III, IV α , IV β , V) ont été comparées à la souche de référence Tohamal par CGH-array en utilisant les microarrays que nous avons développés en 2003 (voir chapitre 2). L'étude a débouché en la détermination de régions génomiques absentes dans certaines souches. Ces régions sont appelées RD ('Region of Differences') et sont au nombre de 10 (RD1 à RD10) pour le panel des souches étudiées dans cet article. Toutes ces RD ont été montrées comme étant flanquées d'IS481 illustrant le mode de perte de matériel génomique par 'looping out' entre régions homologues (McLafferty et al. ; Stibitz, 1998b) et l'importance des IS comme moteur principal de cette dynamique du génome et de l'évolution chez des Bordetelles. Les 9 souches étudiées ont été choisies pour être représentatives géographiquement et temporellement de la collection Française. On a ainsi pu mettre en évidence que les souches circulantes actuellement en France ont moins de matériel génomique que la souche Tohamal. Dans le groupe III les isolats les plus récents ont perdu plus de matériel que les plus anciennes (53kb contre 34kb). Parmi les isolats circulants en France depuis 1990 davantage de matériel génomique a été perdu dans le sous-groupe IV β (57kb) que dans les sous-groupes IV α et V (34kb).

Les régions de différences identifiées contiennent pour certaines des gènes dont la fonction semble essentielle à la vie de la bactérie. Par exemple RD2 qui est absente des isolats circulant actuellement possède des gènes d'acquisition du fer. Il est possible que la restriction de *Bordetella pertussis* à l'hôte humain fait qu'elle ne nécessite plus d'acquérir du fer de différentes sources exogènes et qu'elle peut donc se 'débarrasser' de certains gènes assurant cette fonction. Il est à noter que *Bordetella pertussis* possède plusieurs autres gènes capables d'acquérir le fer (Parkhill et al, 2003).

Pareillement, une autre région de différence, RD10, encode un gène de résistance au cuivre qui est un oligo-élément essentiel mais dont la concentration intracellulaire dans la bactérie doit être finement régulée du fait de sa toxicité même à faible dose. Là aussi il est possible que *Bordetella pertussis* possède d'autres gènes non encore identifiés assurant cette fonction.

Aucun des gènes présents dans les RD identifiées dans cette étude n'est un gène de virulence. Par contre des souches du groupe V montrent des duplications de régions génomiques, en particulier la souche APV23a qui possède une duplication du locus de l'adénylate cyclase (locus *cya*) ainsi que d'un locus impliqué dans la chaîne respiratoire. L'intérêt de ces duplications dans l'adaptation de ces souches reste obscur.

D'un point de vue général, cette publication a permis de confirmer la classification en 5 groupes préalablement établie par PFGE et a défini les caractéristiques génomiques de ces groupes qui peuvent être considérés comme des « genovars » regroupant des souches d'une même espèce sur base de leur contenu en gènes (Porwollik et al., 2004). Cette étude confirme que l'on assiste à une diminution de la diversité génomique par perte d'ORFs au court du temps. Ceci est particulièrement vrai pour le groupe IV β qui est le groupe le plus représenté dans les souches circulantes ces dernières années (depuis 1998). Cette diminution est certainement due à l'adoption par *Bordetella pertussis* d'un mode de vie restreint à un seul hôte, l'humain, ce qui lui permet d'éliminer les gènes non essentiels à ce mode de vie. Ce phénomène a été également constaté chez *Salmonella typhi* et est à opposer à *Salmonella enterica* serovar Thyphimurium (Parkhill et al., 2001).

Un travail similaire a été fait en 2008 pour l'équipe de Roy Gross (Université de Würzburg). Notre équipe a développé pour cette équipe une puce à ADN pour la souche de *Bordetella petrii* et des études de composition génomique par CGH-array ont été réalisées (annexe 3 : Lechner et al., 2008).

Ce travail et la technologie mise au point à l'époque pour obtenir ces données est à rapprocher de travaux beaucoup plus récents dans notre laboratoire. En effet nous utilisons maintenant des approches de séquençage haut débit permettant un séquençage rapide et total de génomes bactériens. Comme à l'époque l'étape d'analyse et d'interprétation des données est cruciale et bien souvent limitante en terme de temps et de qualité d'interprétation. Nous avons développé au laboratoire un pipeline permettant la caractérisation automatique d'un génome bactérien au départ des reads d'un séquençage haut débit. Ce pipeline, appelé MICRA, est en court de publication. Il a été utilisé sur un séquençage de la souche de *Bordetella pertussis* Pillemer P134 qui est une souche utilisée dans les années 1950 comme vaccin.

CHAPITRE 4 : LES PETITS ARN NON-CODANTS CHEZ *BORDETELLA PERTUSSIS*

Annexe 4 : **Hot D**, Slupek S, Wulbrecht B, D'Hondt A, Hubans C, Antoine R, Loch C, Lemoine Y. Detection of small RNAs in *Bordetella pertussis* and identification of a novel repeated genetic element. BMC Genomics. 2011 Apr 27;12(1):207.

La compréhension complète du régulon de BvgAS et plus généralement de la modulation de virulence de *Bordetella pertussis* n'est qu'en partie élucidée. Parmi les questions encore à adresser, on peut se demander si BvgAS est le seul système de régulation de la virulence ou s'il existe des modulateurs auxiliaires agissant sur la modulation de virulence. Est-ce que ces modulateurs auxiliaires fonctionnent indépendamment de BvgA et agissent sur des cibles non-elles même dépendantes de BvgA ou bien s'agit-il de réseaux de régulation interconnectés expliquant l'efficacité d'infection du pathogène ?

De nombreuses études ont maintenant montré l'existence de petits ARNs non-codants qui agissent comme des régulateurs principalement au niveau transcriptionnel ou traductionnel. Certains de ces ARN régulateurs sont des acteurs principaux dans les mécanismes de virulence et de pathogénèse de certaines bactéries (Caldelari et al, 2013).

Les premières étapes de notre réflexion sur ce sujet nous a amené à adresser les questions suivantes : est ce que les Bordetelles, et particulièrement *Bordetella pertussis*, possède des ARN non-codants régulateurs au niveau de son génome, est-ce que ces ARNs sont transcrits et dans quelles conditions et enfin peuvent-ils avoir un lien avec la régulation de la virulence chez cette bactérie ?

Une approche d'analyse bioinformatique de la séquence du génome de *Bordetella pertussis* a été utilisée afin de prédire la présence de petits ARN non-codants dans les régions inter-géniques (IGRs). Cette prédiction s'est faite en utilisant l'algorithme RNAz développé par Washietl et al. en 2005. Il est fondé sur la conservation de structures secondaires prédites entre différents génomes proches, associée à une stabilité thermodynamique significative de cette structure. Les génomes de *Bordetella pertussis* (Tohama I), *bronchiseptica* (RB50), *parapertussis* (12822) et *avium* (N197) ont été comparés dans cette étude et ont permis la prédiction de 23 localisations potentielles sur le génome de *Bordetella pertussis*. Parmi ces prédictions ont été localisé l'ARN 6S qui avait déjà été identifié par homologie (Felden et al., 2001) ainsi que la position de 2 riboswitches déjà référencés à l'époque dans Rfam (<http://rfam.xfam.org/>) : le riboswitch du transcrit de l' amino-méthyl-transférase (BP0195) aussi appelé Glycine-riboswitch (Mandal et al., 2004) et celui du transcrit de ribB (BP0471) qui appartient à la classe des riboswitch-FMN (Flavin MonoNucleotide) (Winkler et al., 2002). La présence de ces riboswitches dans la liste des prédictions n'est pas surprenante car ils présentent les mêmes caractéristiques de conservation de structure secondaire significative.

La transcription au niveau des 20 autres prédictions a été testée *via* la sélection de sondes biotinylées spécifiques du brin positif et du brin négatif des 20 régions

génomiques. Chaque sonde a servi à la révélation de Northern blot obtenus au départ d'ARN total provenant de différentes phases de culture. Une sonde supplémentaire a été sélectionnée pour détecter le petit ARN décrit par Scarlato et al. (Scarlato et al., 1991) positionné en anti-sens du 5'-UTR de l'ARNm de *bvgA*. Cette approche a permis de mettre en évidence 13 régions transcrites en petits ARN non-codants et de confirmer la transcription en anti-sens du 5'UTR de *bvgA*. Ces nouveaux petits ARN non codants ont été dénommés BprA, B, C, ... pour (*Bordetella pertussis* RNA). Le tableau 4.1 résume les caractéristiques de ces nouveaux gènes.

Transcript name	Gene name	Predicted 5' coordinate ^a	Predicted 3' coordinate ^a	Predicted length ^b	Approx. obs. length ^c	sRNA orientation ^d	IGR length ^e	5' gene name and orientation ^f	3' gene name and orientation ^f
TranscriptA1	<i>bprA1</i>	488504	488654	150	250	<	522	BP0475 <	BP0477 >
TranscriptA2	<i>bprA2</i>	488504	488654	150	300	<	522	BP0475 <	BP0477 >
TranscriptB	<i>bprB</i>	1494130	1494207	77	80	>	726	BP1418 <	BP1419 >
TranscriptC	<i>bprC</i>	1968374	1968529	155	190	>	425	BP1878 <	BP1879 >
TranscriptD1	<i>bprD1</i>	2624007	2624157	150	90	<	783	BP2479 <	BP2480 >
TranscriptD2	<i>bprD2</i>	2624007	2624157	150	110	<	783	BP2479 <	BP2480 >
TranscriptE1	<i>bprE1</i>	2699394	2699487	93	70	>	239	BP2546 <	BP2547 >
TranscriptE2	<i>bprE2</i>	2699394	2699487	93	80	>	239	BP2546 <	BP2547 >
TranscriptF1	<i>bprF1</i>	3099570	3099720	150	80	<	256	BP2908 <	BP2909 <
TranscriptF2	<i>bprF2</i>	3099570	3099720	150	150	<	256	BP2908 <	BP2909 <
TranscriptG	<i>bprG</i>	3173584	3173734	150	70	<	305	BP2982 >	BP2983 >
TranscriptH	<i>bprH</i>	3178090	3178331	241	70	>	318	BP2984 >	BP2985 >
TranscriptI	<i>bprI</i>	3263729	3263815	86	450	<	277	BP3061 >	BP3062 >
TranscriptJ	<i>bprJ</i>	3605317	3605397	80	70	>	206	BP3395 <	BP3396 <
TranscriptK	<i>bprK</i>	3619230	3619365	135	200	<	316	BP3410 <	BP3411 >
TranscriptL1	<i>bprL1</i>	3811548	3811699	151	80	<	336	BP3594 <	BP3595 >
TranscriptL2	<i>bprL2</i>	3811548	3811699	151	350	<	336	BP3594 <	BP3595 >
TranscriptL'	<i>bprL'</i>	3811548	3811699	151	80	>	336	BP3594 <	BP3595 >
TranscriptM	<i>bprM</i>	3896371	3896457	86	190	<	281	BP3686 >	BP3687 >
TranscriptM'	<i>bprM'</i>	3896371	3896457	86	190	>	281	BP3686 >	BP3687 >
TranscriptN1	<i>bprN1</i>	3956837	3956934	97	80	<	393	BP3747 >	BP3748 <
TranscriptN2	<i>bprN2</i>	3956837	3956934	97	110	<	393	BP3747 >	BP3748 <

^aGenomic coordinates as predicted by RNAz.

^bIn nucleotides.

^csRNA length, in nucleotides, determined from apparent sizes on Northern blots.

^dGenomic orientation of sRNA transcripts. '>' = on positive strand: the strand given in the GenBank genome database (NC_002929); '<' = on negative strand: the complementary strand.

^eLength, in nucleotides, of the IGRs containing sRNA genes (according to the annotation).

^fGenomic orientation of flanking mRNAs. '>' = on positive strand: the strand given in the GenBank genome database (NC_002929); '<' = on negative strand: the complementary strand.

Tableau 4.1

La transcription de ces Bpr a été analysée d'après les profils obtenus sur Northern blot (voir papier annexe 4). L'expression de ces transcrits en condition de modulation phénotypique (en présence de MgSO₄ ou d'Acide Nicotinique dans le milieu de culture) ou dans un mutant avirulent (suite à la délétion de *BvgAS*) a permis de déterminer que *bprJ2* étaient sous le contrôle de *BvgA*. Ce petit ARN fait actuellement l'objet d'études plus approfondies dans notre laboratoire (voir chapitre 7).

CHAPITRE 5 : LE TRANSCRIPTOME PRIMAIRE DE *BORDETELLA* *PERTUSSIS*

Travaux en cours

L'étude de la régulation transcriptomique d'une bactérie pathogène et de l'importance de cette régulation sur sa virulence passe ces dernières années par l'établissement de la composition et de la cartographie du transcriptome complet de la bactérie en question. Ce niveau de détails peut être atteint grâce au cumul des efforts de séquençage des génomes entiers des années 1990 et des annotations qui ont découlé de ces séquençages ainsi que des applications de RNA-seq issues des séquenceurs à haut-débit depuis les années 2010. Les résultats obtenus par ces approches ont montré une complexité de transcriptome insoupçonnée dans différentes familles bactériennes (Croucher *et al.*, 2010 ; van Vliet *et al.*, 2010 ; Dugar *et al.*, 2013).

Nous avons utilisées ces approches de séquençage haut débit sur le transcriptome de *Bordetella pertussis* souche Tohamal cultivé en condition standard (milieu Stainer-Scholte (SS) à 37°C jusqu'à la phase exponentielle de croissance). Deux approches indépendantes de séquençage ont été combinées afin de maximiser les données concernant l'architecture du transcriptome. Un RNA-seq profond sur une librairie d'ADNc orientés a tout d'abord été obtenu en utilisant le protocole TruSeq Stranded mRNA et le séquenceur HiSeq d'Illumina sur un ARN préalablement déplété en ARNr. Les 'reads' de séquençage obtenus ont été filtrés en fonction de leur qualité et pour éliminer ceux s'alignant sur les séquences des clusters d'ARNr. Après filtre, 26 millions de 'reads' ont pu être alignés sur le génome de référence de Tohamal (NC_002929) en utilisant Bowtie2 (Langmead *et al.*, 2012). Le niveau d'expression des gènes annotés a ensuite été évalué en utilisant EDGE-pro (Estimated Degree of Gene expression for Prokaryote) (Magoc *et al.*, 2013).

Nous avons également utilisé l'approche de 'differential RNA-seq' (dRNA-seq) décrite par Sharma *et al.* (Sharma *et al.*, 2010) dans le but de détecter les sites de départ de transcription (en anglais : Transcription Starting Sites : TSS). Le kit Ion Total RNA-seq v2 (Life Technologies) a été utilisé sur de l'ARN de *Bordetella pertussis* préalablement traité ou non-traité à la 5'terminal phospho-dépendante exonucléase (TEX). Les deux librairies ainsi obtenues, traitée et non-traitée, ont été séquencées sur l'Ion Torrent PGM (Life Technologies). Le protocole complet a été répété 3 fois afin d'obtenir 3 réplicats biologiques. Cette approche permet de détecter les transcrits primaires résultant d'une initiation de transcription par la RNA-polymérase (Wek *et al.*, 1986). Un total de 33 millions de 'reads' a été généré pour cette expérience (17,4 millions pour les échantillons traités à la TEX et 15,7 millions pour les non-traités). Les TSS résultant des dRNA-seq ont été déduits en utilisant l'algorithme TSSAR (Amman *et al.*, 2014) suivi d'une curation manuelle des données.

Ces approches ont permis d'établir une série de résultats résumés sous forme d'un navigateur web présentant les résultats obtenus en parallèle de l'annotation et des coordonnées du génome. L'ensemble des résultats concourent à la définition du transcriptome de *Bordetella pertussis*. La figure 5.1 est un ensemble de captures d'écran représentant les résultats visibles au niveau du navigateur web.

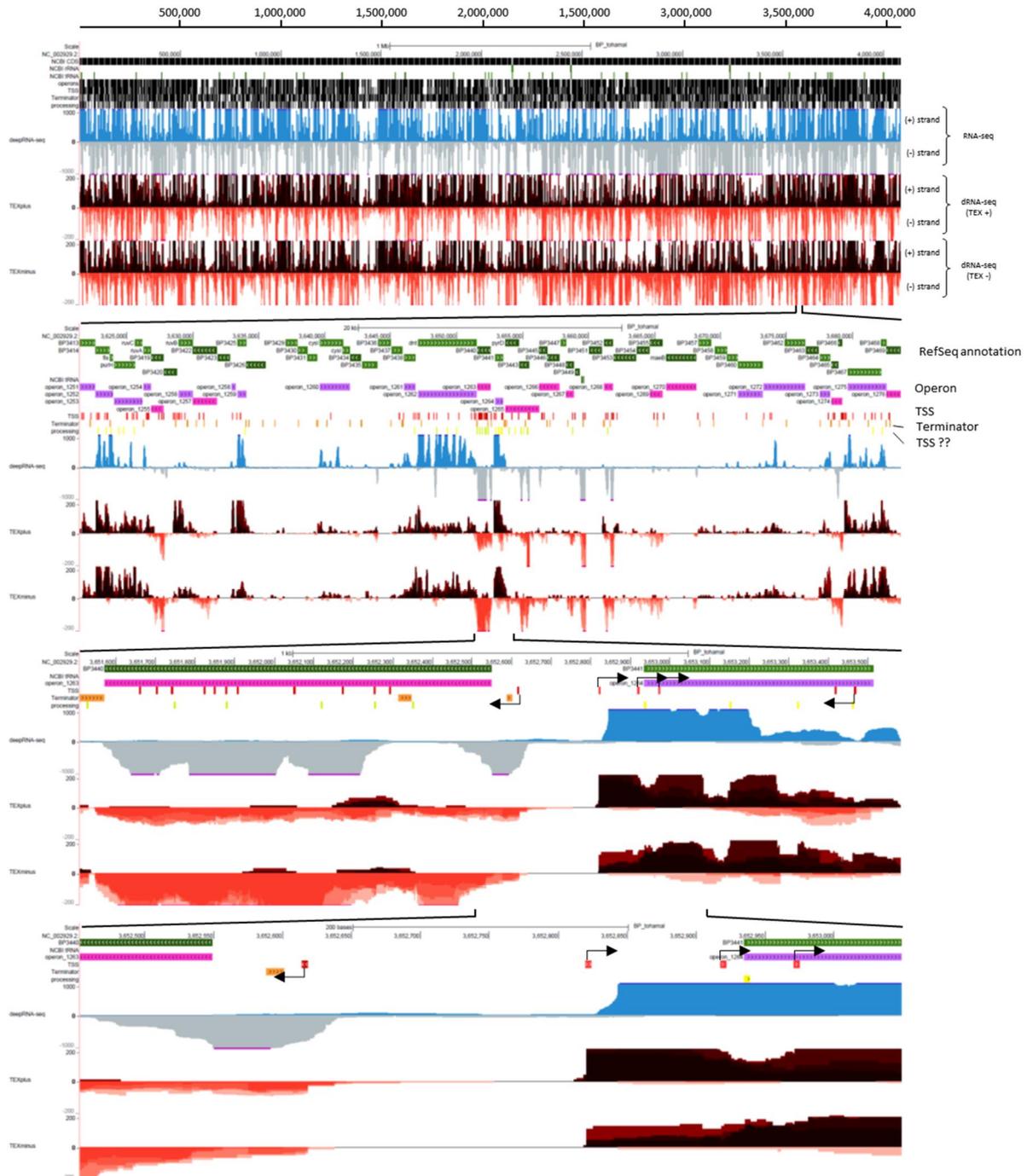


Figure 5.1 : Captures d'écrans à 4 niveau de zoom différents du « Hub » intégrant l'ensemble des données et informations issues de l'étude du transcriptome primaire de *Bordetella pertussis*. Les données en bleu et gris correspondent au graphe de profondeur du RNA-seq Illumina. Les données en rouge et marron correspondent aux graphes de profondeur des dRNA-seq (3 réplicats représentés par nuance de rouge et de marron). Les flèches en vert représentent les gènes annotés (d'après RefSeq NC_002929). Les flèches en mauve représentent les prédictions d'opérons. Les flèches en rouge représentent les TSS. Les flèches en orange représentent les prédictions de terminateurs.

L'analyse de ces données nous a permis de définir ou redéfinir un ensemble de caractéristiques du transcriptome de *Bordetella pertussis* :

- Organisation des unités de transcription pour 2499 des 3871 CDS codants, formant 1439 structures opéroniques (996 monocistroniques et 443 polycistroniques),
- Définition des +1 de transcription pour 614 gènes ou opérons incluant 35 transcrits sans 'leader' ('leaderless mRNA'),
- Prédiction de 840 zones génomiques réparties sur l'ensemble du génome pour des candidats de petits ARN non codants,
- La détection de 7 riboswitches dont 3 ont été validés.

L'ensemble de ces caractéristiques sont illustrées et approfondies ci-dessous :

5.1 Définition des structures opéroniques chez *Bordetella pertussis*

Les données obtenues du RNA-seq profond ont été utilisées pour caractériser la structure et l'organisation des transcrits en regard de l'annotation du génome. Une organisation de transcription a pu être déduite de cette analyse pour 2499 des 3871 CDS annotées (en comptant les IS). Le premier niveau d'analyse a été consacré aux structures opéroniques afin de déduire les transcrits mono et poly-cistroniques. Un script a spécifiquement été écrit par notre collaborateur Fabian Amman du TBI-ViennaRNA (non publié) afin de déduire les structures opéroniques des résultats d'alignement des reads. Cette analyse a permis de déduire 1439 structures de transcrits, 996 d'entre elles sont monocistronique et 443 polycistroniques. Une comparaison avec une prédiction purement algorithmique telle que celle proposée par la base de données DOOR2 (Database of prokaryotic OperOn) (Mao et al., 2009) révèle 859 structures identiques (617 monocistroniques et 242 polycistroniques).

5.2 Les +1 de transcription : 'Transcription Starting Starts' (TSS)

L'analyse de nos données de 'differential RNA-seq' a permis de définir 2722 sites de potentiels +1 de transcription qui ont été positionnés par rapport à l'annotation du génome. En fonction de leur position relative par rapport aux ORF et de la même manière que précédemment définie (Sharma *et al.*, 2010), les TSS ont été classés comme Primaires (P) ou Secondaire (S) lorsqu'ils sont en amont d'une ORF, comme Internes (I) lorsqu'ils sont à l'intérieur d'un ORF et dans la même orientation ou comme Antisens (A) lorsqu'ils sont en orientations opposées. Enfin les TSS situés dans les régions intergénomiques sont définis comme Orphelins (O).

Si on considère uniquement les gènes codants, 614 TSS primaires ont été déterminés et peuvent être considérés comme le +1 de transcription de ces gènes. Pour 99 de ces ORF un ou plusieurs TSS secondaires ont été détectés. Seulement 24 positions de +1 de transcription sont décrites à ce jour dans la littérature pour le génome de *Bordetella pertussis* Tohamal. Nous avons comparé les positions des +1 publiés à la liste de nos TSS. Six de ces positions correspondent à des gènes non transcrits dans nos conditions d'étude et ne peuvent donc pas donner de détection de TSS. Des 18 positions restantes, 7 TSS ont été détectés à l'exacte position publiée ou à une position très proche de la position publiées (à +/- 1 ; 2 ou 3 nucléotides). Pour les autres positions publiées il n'y a pas de TSS détectés (principalement par absence de différence entre les bibliothèques traitées et non-traitées à la TEX) montrant que la méthode utilisée par 'differential RNA-seq' permet une détection efficace mais pas complète des départs de transcription.

Parmi les TSS primaires, 35 ont été détectés à l'exacte position du codon 'start' de l'ORF du gène annoté. Les transcrits de ces gènes sont donc très probablement sans 5'UTR (: 'leaderless'). Un tiers de ces gènes sont de fonction inconnue mais une large proportion des gènes de fonction connue (24%) sont classés comme 'Régulateurs'. Cette proportion a été calculée comme très significativement enrichie par un test statistique hypergéométrique (p-Val = 0.0046).

5.3 Prédiction et détection de sRNA

La détection des positions des TSS sur un génome bactérien par l'approche de 'differential RNA-seq' est généralement également exploitée pour déterminer les positions de potentiels nouveaux transcrits et particulièrement celles des petits ARN non codant. Nous avons entamé l'établissement d'un répertoire de ces petits ARN chez *Bordetella pertussis* (voir chapitre 4, Hot et al., 2011) avec une approche assez stringente basée uniquement sur une prédiction algorithmique. La liste des TSS obtenue ici par 'differential RNA-seq' nous permet de compléter ce répertoire. Tous les TSS classés comme Orphelins ou Antisens (n=392) sont considérés comme de potentiels sites de départ de nouveaux candidats de sRNA. Une rapide comparaison des TSS et de la liste publiée en 2011 montre que seulement la moitié des sRNA validés en 2011 sont retrouvés dans la liste des 392 potentiels nouveaux sites. Afin de produire un répertoire le plus complet possible, nous avons donc décidé de cumuler ces prédictions obtenues par une approche de séquençage à des prédictions obtenues par des approches *in silico*. Trois nouveaux algorithmes basés sur des critères de sélection différents ont été utilisés : SIPHT (sRNA Identification Protocol using High-throughput Technologies) qui se fonde sur de la conservation de la séquence en amont

de terminateurs (Livny et al, 2012) ; NAPP (Nucleic Acids Phylogenetic Profiling) qui utilise une comparaison de séquences des régions intergéniques (Ott et al, 2012) et le programme RNAz qui est basé sur de la prédiction de structure stable et significativement conservées. Ces 3 approches *in silico* ont donné un total de 509 candidats. Les 3 algorithmes semblent assez complémentaires car seulement 15 prédictions sont communes aux trois algorithmes et seulement 84 sont trouvés par 2 des 3 algorithmes. De plus, seulement 41 prédictions sont communes entre les 392 obtenues par 'differential RNA-seq' et les 509 obtenues par séquençage.

Afin d'évaluer la valeur des prédictions des différentes approches (par séquençage ou par prédiction *in silico*), 58 d'entre-elles ont été validées par détection de transcrits sur Northern blot. Les approches *in silico* donnent un taux de vrais positifs de 82% (avec 70% de précision et 82% de sensibilité) contre 49% de vrais positifs (avec 100% de précision et 49% de sensibilité) pour l'approche en séquençage. Si l'on englobe les résultats des 2 approches, en considérant comme vrai positif un candidat validé qui est prédit par l'une ou l'autre des approches, on atteint 92% de vrais positifs (avec 72% de précision et 92% de sensibilité). Cela signifie que le cumul de 2 approches, l'une se basant sur les connaissances que l'on a des petits ARN non codant et leurs caractéristiques et l'autre sans *a priori*, permet de virtuellement déterminer la position de 9 sRNA sur 10 en faisant moins de 30% de fausses prédictions.

Une attention particulière a été portée sur les prédictions de petits ARN situées à proximité ou en antisens des gènes de virulence de la bactérie. En dehors du petit ARN BprC, situé en antisens du 5'UTR de *bvgA* et dont l'étude fera l'objet du chapitre 7, notre approche a permis de déterminer 11 prédictions proche de gènes de virulence. Quatre en antisens de gènes d'adhésines (*fhaB*, *fim2*, *fim3*, *fimB*), 3 en antisens de toxine (*cyaA*, *dnt* (région interne), *dnt* (en 3')) et 5 en antisens de gènes de systèmes de sécrétions (*cyaD*, *ptlC*, *ptlH*, *bcrH2*, *bcsN*). Pour l'instant 4 de ces petits ARN ont pu être validés par détection d'un transcrit en Northern blot (*fhaB*, *fim2*, *cyaA*, *dnt* (en 3')). Quatre autres se sont révélés négatifs à la validation en Northern blot (*fim3*, *dnt* (région interne), *ptlC* et *ptlH*) et ceci malgré une détection assez nette en séquençage. Une seconde sonde biotinylée va être testée afin de s'assurer de cette absence d'expression. Enfin 4 prédictions restent encore à tester en Northern blot (*fimB*, *cyaD*, *bcrH2*, *bcsN*). Le mode de régulation de ces sRNA antisens et leurs potentielles fonctions sont en cours d'étude.

5.4 Les riboswitches

Parmi les éléments remarquables et d'intérêt dans le transcriptome d'une bactérie on distingue bien souvent les transcrits présentant un riboswitch. Il s'agit d'une structure

située dans la partie du 5'UTR d'un ARNm et ayant une fonction régulatrice. Cette structure peut être fixée par un ligand (souvent une petite molécule intracellulaire) et modifier l'expression du gène en question par activation ou blocage de la traduction. Plusieurs mécanismes d'activation et de traduction existent. Une recherche sur ces structures a été initiée avec Rudy Antoine (équipe BRIPT-CIIL) en utilisant le programme Infernal (E. P. Nawrocki and S. R. Eddy) qui est utilisé pour la base de données Rfam (<http://rfam.xfam.org/>). Une liste de 7 riboswitch différents a pu être établie, le tableau 5.1 reprend les caractéristiques de ces éléments.

target	name	accession	from	to	strand	E-value
SAM_alpha	RF00521	BP0047_promoter_47985_48208	48078	48158	+	3.7e-08
Glycine	RF00504	BP0195_promoter_197008_197462	197088	197180	+	2.8e-11
Glycine	RF00504	BP0195_promoter_197008_197462	197202	197321	+	7.3e-12
TPP	RF00059	BP0206_promoter_209453_209831	209596	209707	+	2.4e-07
FMN	RF00050	BP0471_promoter_481772_482205	481881	482034	+	3.1e-29
SAH_riboswitch	RF01057	BP3068_promoter_3270966_3271305	3271127	3271247	-	4.1e-09
YybP-ykoY (manganese)	RF00080	BP3410_promoter_3619131_3619507	3619234	3619427	-	9.6e-15
Cobalamin	RF00174	BP3595_promoter_3811442_3811838	3811547	3811706	+	1.1e-11

Tableau 5.1 : Prédiction des riboswitches en utilisant Infernal

La validité de 5 de ces riboswitches (Glycine, FMN, SAH, TPP, SAM) a été éprouvée par l'utilisation de couples d'amorces de RT-PCR situés dans le 5'UTR, dans la première partie de l'ORF et à cheval entre le 5'UTR et le début de l'ORF. Deux autres riboswitches (YybP-ykoY et Cobalamine) sont en cours d'évaluation mais correspondent probablement à des gènes non transcrits dans les conditions standards de culture.

CHAPITRE 6 : ETUDE DES FONCTIONS DU FACTEUR HFQ CHEZ *BORDETELLA PERTUSSIS*

Annexe 5 : Bibova I, Skopova K, Masin J, Cerny O, **Hot D**, Sebo P, Vecerek B. The RNA chaperone Hfq is required for virulence of *Bordetella pertussis*. Infect Immun. 2013 Nov;81(11):4081-90.

Annexe 6 : **Hot D***, Bibova I*, Keidel K, Amman F, Slupek S, Cerny O, Gross R, Vecerek B. Transcriptional profiling of *Bordetella pertussis* reveals requirement of RNA chaperone Hfq for Type III secretion system functionality. RNA Biol. 2015 *: co-first authors

La protéine Hfq, initialement identifiée comme un facteur nécessaire à la réplication du bactériophage Q β chez *E. coli* (Franze et al., 1968), est en fait une chaperonne d'ARN impliquée entre autre dans la reconnaissance de certains petits ARN non codants bactériens et de leurs ARNm cibles. Elle agit en fixant à la fois le petit ARN et son ARNm cible (Mikulecky et al., 2004 ; Link et al., 2009), favorisant ainsi l'interaction par modifications conformationnelles et changement de la stabilité des 2 acteurs (Geissmann et al, 2004 ; Moll et al., 2003 ; Sledjeski et al., 2001 ; Afonyushkin et al., 2005). Plusieurs études de mutants, dont le gène *hfq* a été délété, ont montré un effet pléiotropique comme par exemple chez *E. coli* où l'on observe une diminution de la vitesse de croissance et une augmentation de la sensibilité au stress (UV, osmolarité) (Tsui et al., 1994) montrant le rôle général joué par Hfq dans la physiologie de la bactérie. Chez diverses bactéries, les petits ARNs non-codants associés à Hfq ont été montrés comme étant impliqués dans une variété de processus cellulaires tels que la répression catabolique, l'homéostasie du fer, l'intégrité de l'enveloppe cellulaire et la pathogénèse (Masse et al., 2002 ; Zhang et al., 1998 ; Vecerek et al., 2007 ; Udekwu et al., 2005 ; Sonnleitner et al., 2009). Plus spécifiquement Hfq est apparue comme essentielle à l'aptitude physiologique et à la virulence d'une large gamme de bactéries pathogènes (Chao et al., 2010).

Ces observations et plus particulièrement le lien avéré entre Hfq et la virulence de plusieurs pathogènes nous a incité à étudier ce facteur chez *Bordetella pertussis* (Annexe 5 : Bibova et al, 2013). Ce travail a été fait en collaboration avec Branislav Vecerek de l'Institut de Microbiologie de Prague (ASCR, République Tchèque) qui est spécialiste d'Hfq et qui travaille également sur la problématique des petits ARNs chez les Bordetelles.

Un mutant délété au niveau de la phase de lecture ouverte du gène *hfq* a été généré par clonage en prenant soin de conserver les régions du promoteur et du terminateur de transcription. L'effet de l'absence d'Hfq sur la physiologie de la cellule a tout d'abord été évalué au niveau de la croissance au cours du temps dans un milieu SS standard à 37°C. La croissance du mutant a été comparée à celle d'une souche sauvage pendant 34h de culture. Il apparaît que le mutant est nettement impacté par la délétion avec un temps de doublement qui passe de ~4h à ~6h entre la souche sauvage et le mutant et un maximum de densité optique ~2x moins important. Une série de facteurs de virulence a été étudié dans le mutant Δhfq et comparé à la souche sauvage. Il apparaît que l'adénylate cyclase (ACT) présente une diminution à la fois du taux de production de la protéine et de son taux de sécrétion dans le milieu extérieur. La toxine pertussique a également été étudiée en évaluant la production de la sous-unité S1. Il apparaît que la production de la toxine pertussique ne semble pas affectée par

l'absence du facteur Hfq par contre sa sécrétion est nettement diminuée. De la même façon la sécrétion mais pas la production de l'hémagglutinine filamenteuse (FHA) semble affectée à la baisse par l'absence d'Hfq dans la cellule lors de la phase exponentielle de culture.

Cette modification dans la production et/ou sécrétion de facteurs de virulence nous a encouragé à regarder l'effet de l'absence d'Hfq sur la virulence chez la souris. En particulier la diminution d'ACT produite et sécrétée devrait altérer la capacité de *Bordetella pertussis* à modifier la voie de signalisation par l'AMP cyclique des cellules immunitaires qui habituellement inhibe les capacités chimiotactiques, phagocytiques et le stress oxydatif associés à l'infection assurant la persistance de la bactérie chez l'hôte infecté. Des souris ont été inoculées à $1,5 \times 10^8$ UFC/souris (Unité Formant une Colonie par millilitre) dans le nez avec la souche sauvage, la souche Δhfq ou la souche Δhfq complémentée. Une mortalité de 100% a été observée en 4 jours chez les souris inoculée avec la souche sauvage alors qu'aucune mortalité n'a été observée pour le mutant Δhfq . Une courbe similaire de mortalité n'a pu être obtenue pour le mutant qu'à une dose de $1,5 \times 10^9$ UFC/souris (voir figure 6.1-A). La dose létale à 50% (DL₅₀) issue de ces expériences donne des valeurs de 6×10^7 , 7×10^7 et 8×10^8 UFC/souris pour la souche sauvage, la souche complémentée et le mutant respectivement.

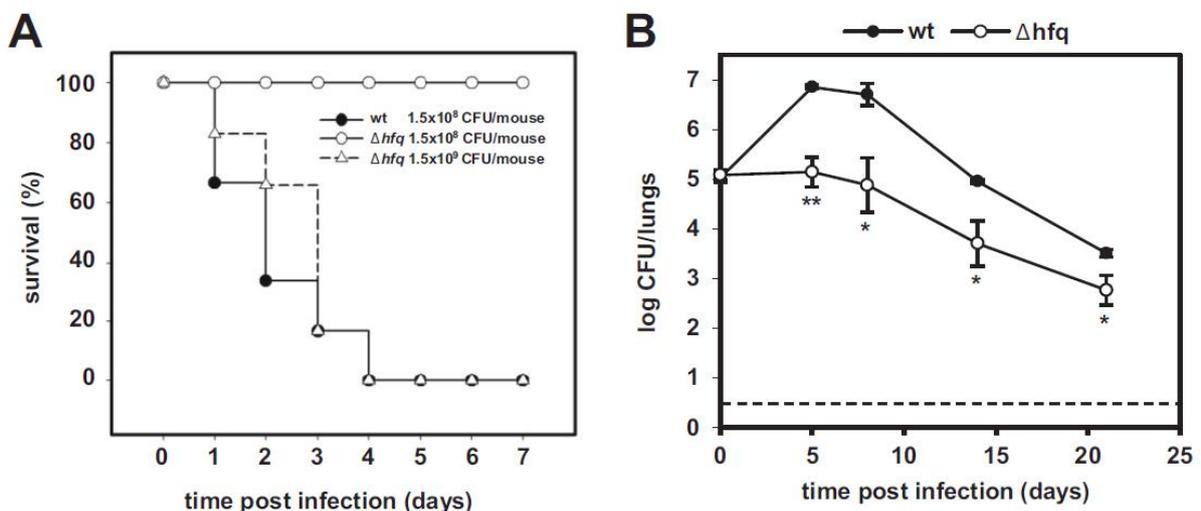


Figure 6.1 : Effet de la délétion *dhfq* sur la virulence de *Bordetella pertussis*. (A) Courbes de survie des souris post-infection pour la souche sauvage (cercles fermés), chez le mutant Δhfq à $1,5 \times 10^8$ CFU/souris (cercle ouverts) et le mutant Δhfq à $1,5 \times 10^9$ CFU/souris (triangles). (B) Colonisation des poumons après inoculation de la souche sauvage (cercles fermés) et de la souche mutante (cercles ouverts). D'après Bibova et al., 2013.

Dans un second temps la capacité de colonisation du poumon chez la souris par le mutant Δhfq a été évaluée. Une dose non létale de 2×10^5 UFC/souris a été utilisée à l'infection et le nombre de UFC par poumon a été mesuré jusqu'à 20 jours post inoculation. La souche sauvage prolifère dans le poumon pour atteindre 7×10^6

CFU/poumon à 5 et 8 jours post-infection alors que dans le même temps le mutant Δhfq n'est pas capable de proliférer et de coloniser le poumon (voir figure 6.1-B). Cette conclusion est également valable avec une dose d'inoculation à 2×10^6 CFU/souris qui compense le temps de génération plus long du mutant. Le défaut de colonisation observé chez le mutant est donc probablement dû à la réduction d'expression des facteurs de virulence plutôt qu'à la réduction du taux de croissance. Des expériences complémentaires d'infection mixte avec la souche sauvage et le mutant en compétition ont permis de montrer une plus faible capacité de persistance du mutant lors de l'infection.

Cette première étude sur le facteur Hfq de *Bordetella pertussis* nous a permis de mettre en évidence son implication dans des mécanismes de régulation impliqués dans la production de facteurs de virulence et dans la pathogénicité de la bactérie. Nous avons décidé de poursuivre cette investigation en étudiant le régulon complet de facteur Hfq. Le microarray développé dans notre labo et décrit en chapitre 2 a été engagé dans une étude du transcriptome complet du mutant Δhfq en le comparant au transcriptome de la souche sauvage (annexe 6). Plusieurs centaines de gènes ont été mesurés comme étant modulés (soit surexprimés soit en sous-exprimés) dans le mutant par rapport à la souche sauvage et ceci en phase exponentielle comme en phase stationnaire. Les gènes obtenus ont été regroupés par classes fonctionnelles. Il apparaît que la classe « Energy metabolism » en phase exponentielle et les classes « Pathogenicity and adaptation » et « Information transfert » en phase stationnaire sont significativement enrichies en représentativité. Parmi les gènes présentant une modulation plusieurs facteurs de virulence sont nettement sous-exprimés dans le mutant par rapport à la souche sauvage en phase exponentiel comme en phase stationnaire. On trouve dans cette liste le gène de l'autotransporteur *vag8*, le gène de résistance au sérum *brkA*, le facteur de colonisation trachéal *tcfA* et enfin une bonne partie des clusters de gènes impliqués dans le système de sécrétion de type 3 (T3SS). Les gènes de l'opéron de la toxine pertussique et de son système de sécrétion sont également modulés mais contrairement aux autres facteurs, ils apparaissent surexprimés dans le mutant.

Une étude approfondie de la régulation du T3SS a été ensuite initiée. Notre étude transcriptomique a montré que les gènes du locus *bsc* ainsi que les gènes *bsp22*, *bopD* et *bopN* (codant pour le 'tip', le translocateur formant le pore et l'effecteur du T3SS respectivement) sont sous-exprimés dans le mutant. Il est important de noter que le locus *btr*, qui comprend les gènes régulateurs du T3SS, n'est pas affecté par l'absence d'Hfq. Par contre les gènes *bteA* et *btcA*, qui codent pour un autre effecteur et sa

chaperonne respectivement, sont eux sous exprimés tout comme le locus *bsc* alors qu'ils ne sont pas localisés dans ce locus.

La souche de *Bordetella pertussis* adaptées ne produisent pas les protéines du T3SS mais il a été montré qu'une telle souche recouvrait la production de Bsp22 après passage chez la souris et ceci pendant plusieurs passages. Nous avons pu montrer que l'activation de la production de ce 'tip' (et donc la fonction du T3SS) lors du passage chez la souris était Hfq dépendante. Ce même type d'observation a été fait lors d'une infection de macrophage en culture. La production de Bsp22 est observée dans la souche sauvage après seulement 3h de contact avec les macrophages et cette réactivation d'expression du T3SS n'est pas observée chez le mutant Δhfq démontrant que Hfq est nécessaire à la synthèse des substrats du T3SS *in vivo*.

CHAPITRE 7 : ETUDE DU PETIT ARN NON-CODANT, BPRC.
IMPLICATION DANS LA REGULATION DE LA VIRULENCE DE
BORDETELLA PERTUSSIS

L'objectif principal de notre projet sur les petits ARNs non-codants est de caractériser et comprendre les mécanismes moléculaires de régulation dans lesquels ils sont impliqués. Nos efforts sont essentiellement tournés vers la compréhension des mécanismes de virulence de *Bordetella pertussis* et sa capacité à contrecarrer les défenses immunitaires. Le choix des petits ARNs non-codants à investiguer se porte donc sur des candidats ayant des implications potentielles dans la régulation des mécanismes de virulence : la reconnaissance et l'adhésion à des cellules cibles, le détournement de la réponse immunitaire, la production de toxines ou plus généralement de facteurs de virulence.

Un des premiers choix d'investigation a été le petit ARN non-codant BprC. Ce petit ARN a été caractérisé lors de notre première étude en 2011 (Hot et al, 2011) mais il avait été détecté fortuitement en 1998 par Scarlato (Scarlato et al, 1991) lors d'une étude du promoteur du système à deux composants *bvgAS*, régulateur majeur de la virulence chez *Bordetella pertussis*. Le choix de ce petit ARN résulte principalement de sa position en antisens de la région 5' non codante (5' UTR) de l'ARNm de *bvgA*. En effet nous avons pu caractériser à la fois par 5' RACE ('Rapid Amplification of cDNA Ends') et par 'differential RNA-seq' la position probable de départ du transcrit de *bprC*. Il est positionné à 14 nucléotides en amont du codon start de BvgA, dans l'orientation opposée et s'étend sur un peu moins de 200 nucléotides d'après sa taille apparente en Northern blot. Il couvre donc en antisens la totalité du 5'UTR de *bvgA* démarré de ses promoteurs 1 ou 2 et des 3/4 du 5'UTR de *bvgA* démarré à son promoteur 3 (voir figure 7.1).

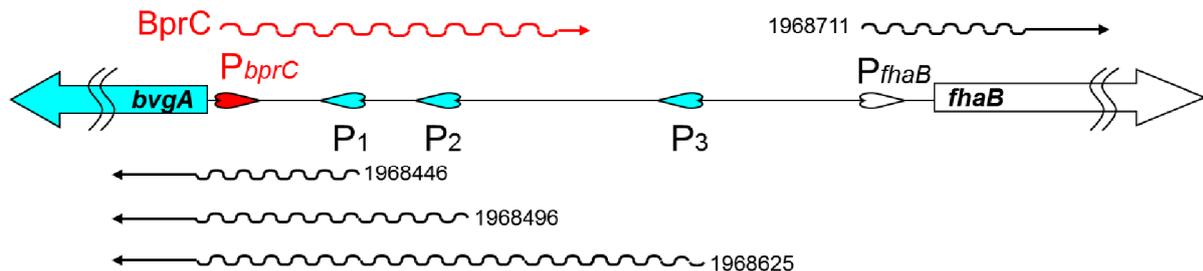


Figure 7.1 : Organisation des départs de transcription dans la région intergénique entre *bvgA* et *fhaB*. Les ORF sont représentés par de larges flèches et les ARNs par des flèches en vagues. Les chiffres correspondent aux coordonnées des départs prédits.

L'expression de BprC a été évaluée dans différentes conditions de culture. Nous avons tout d'abord étudié son expression pendant la phase de culture en prélevant toutes les 0,5 unités de DO_{600nm} entre 0,5 et 4,0. Il apparaît que BprC est transcrit très rapidement (dès $OD = 0,5$) et atteint un maximum à $OD = 1,0$ pour ensuite rester constant. Son expression a également été étudiée dans les différentes phases de virulence de la bactérie. En phase *bvg-*, dans un mutant $\Delta bvgA$ ou en présence de $MgSO_4$ le transcrite est absent ou très faiblement détecté en Northern blot (voir figure 7.2-A). L'expression de BprC a ensuite été étudiée lors du shift de la phase *bvg-* à la phase *bvg+* par le suivie d'une culture juste après passage de 20°C (*bvg-*) à 37°C (*bvg+*). BprC est absent à 20°C et son expression est détectable seulement 10 minutes après le shift (voir figure 7.2-B).

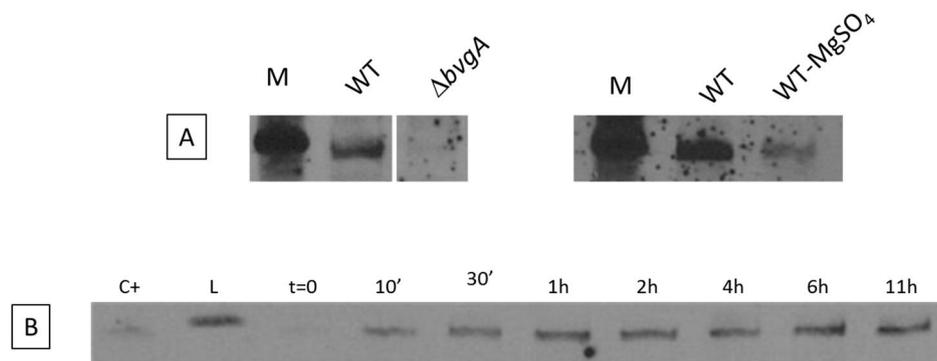


Figure 7.2 : Etude de l'expression de BprC en phase *bvg-* et *bvg+* et lors du shift d'une phase à l'autre. (A) expression en mutant $\Delta bvgA$ et en présence de $MgSO_4$. (B) Shift de culture de 20°C (t=0) à 37°C. Les temps indiqués correspondent au temps de récolte après le shift.

BprC a donc un profil d'expression de *vag* ('virulence activated gene'), il n'est pas exprimé en phase *bvg-* et est rapidement exprimé lorsque la bactérie passe en phase *bvg+*.

Nous avons ensuite muté le promoteur potentiel de *bprC* en introduisant des mutations dans les boîtes -10 et -35 afin d'éliminer l'expression de *bprC* et ainsi pouvoir étudier sa fonction. Les mutations créées dans le promoteur de *bprC* ont dû être faites en tenant compte des codons du début de l'ORF de *bvgA* afin de ne pas modifier la séquence protéique de BvgA. Ces mutations ont provoqué non pas l'extinction complète de l'expression de BprC mais ont permis de diminuer de 80% cette expression (calculée par RT-qPCR) ce qui est suffisant pour mesurer l'effet de cette réduction. Les caractéristiques du mutant obtenu ont tout d'abord été évaluées en termes de comportement en culture. La courbe de croissance du mutant est fortement perturbée avec un décrochement par rapport à une souche sauvage qui s'opère dès la phase d'accélération et qui se prolonge tout au long de la culture pour finir à un

plateau ~30% plus bas (voir figure 7.3-A). La mise en culture sur de l'agar au sang a également permis de mettre en évidence que l'absence de BprC résulte en une perte de l'activité hémolytique (voir figure 7.3-B). Cette perte d'activité hémolytique est rétablie par complémentation du mutant ce qui prouve l'implication de ce petit ARN dans ce phénotype de virulence (voir figure 7.3-C).

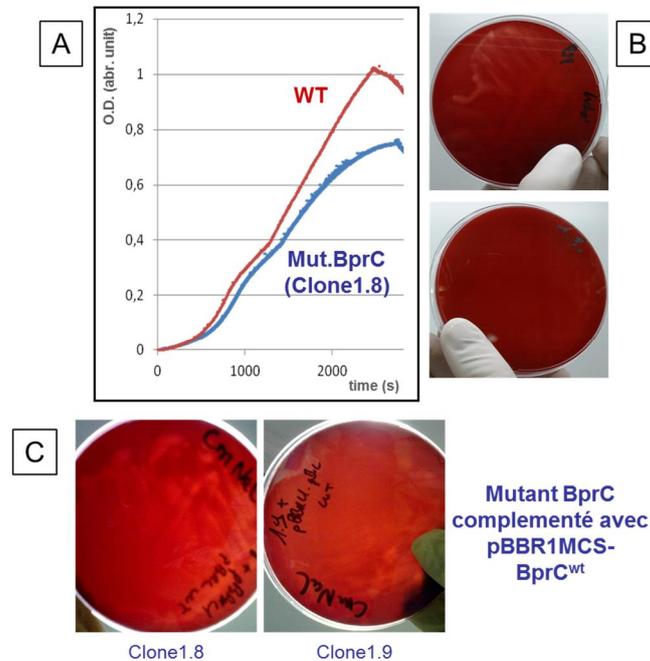


Figure 7.3 : Étude des caractéristiques du mutant BprC en culture. (A) Mesure de la densité optique (600 nm) au cours du temps du mutant BprC (courbe en bleue, Mut.BprC) et de la souche sauvage (courbe en rouge, WT). (B) Mise en évidence de l'activité hémolytique par mise en culture sur agar au sang. (C) mise en culture sur agar au sang du mutant complétement avec la version sauvage du promoteur de *bprC*.

Le mutant *bprC* a ensuite été engagé dans une étude transcriptomique en utilisant le microarray mise au point par notre équipe (voir chapitre 2) afin de déterminer les gènes affectés par la diminution de BprC dans la cellule et de comprendre la fonction du petit ARN. Le mutant *bprC* présente un profil transcriptomique très perturbé par rapport à la souche sauvage avec 83 gènes sous-exprimés et 12 gènes sur exprimés. Très vite il est apparu que ce régulon ressemblait au régulon de BvgA. Une comparaison précise des deux régulons (celui du mutant BprC et celui de $\Delta bvgA$) a été faite graphiquement en représentant les ratios de modulation (en log Fold-Change) du mutant BprC en fonction des ratios du mutant $\Delta bvgA$ (figure 7.4). Cette comparaison a permis d'établir que la grande majorité des *vag* du régulon de BvgA sont également sous exprimés chez le mutant *bprC* (figure 7.4 cercle bleu) mais un certain nombre restent non modulés chez le mutant *bprC* (figure 7.4 cercle vert), c'est le cas de *bvgA* lui-même qui est connu pour être un *vag* auto-régulé. Pareillement, les *vrg* ne sont pas modulés chez le mutant *bprC* (figure 7.4 cercle rouge). Il semble donc qu'une partie,

mais pas la totalité, du régulon de BvgA soit affecté par la diminution de BprC dans la cellule. Cette observation peut signifier que BprC agit, non pas directement sur ces gènes, mais *via* une action sur concentration de BvgA dans la cellule. Afin de déterminer si certains gènes sont affectés directement par la diminution de BprC, et non pas uniquement *via* BvgA, un double mutant $bprC + \Delta bvgA$ a été comparé à la modulation obtenu dans le mutant $\Delta bvgA$ uniquement. Les valeurs de modulation entre ces deux comparaisons s'alignent parfaitement avec la médiatrice, les deux comparaisons donnent donc exactement le même profil de modulation. Le fait de faire un double mutant BvgA et BprC n'a donc pas plus d'effet de modulation que le simple mutant $\Delta bvgA$. Cette observation confirme que BprC agit sur son régulon *via* une action sur BvgA.

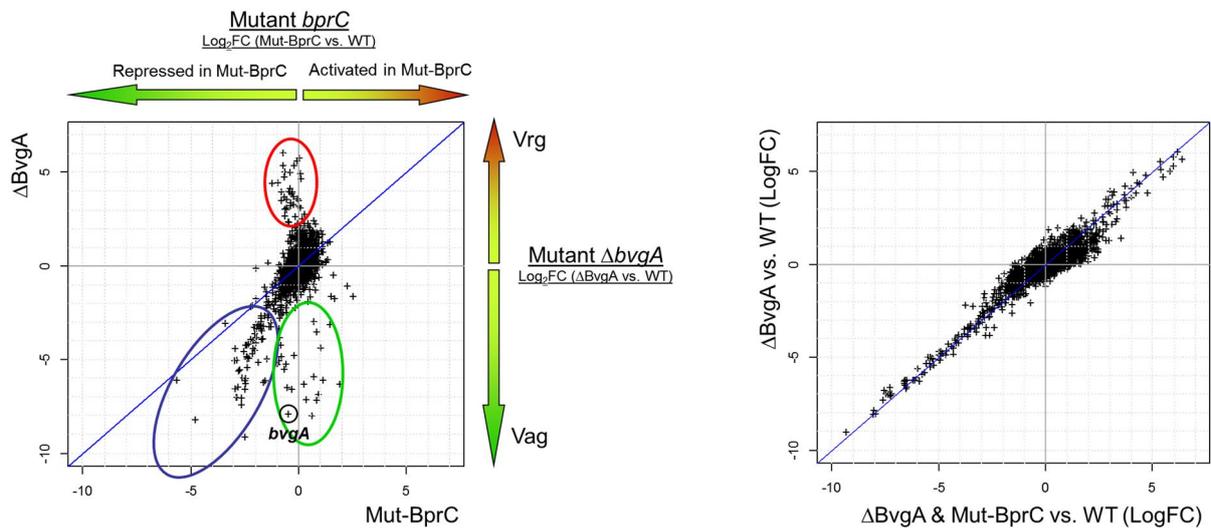


Figure 7.4 : Comparaison des ratios de modulation (en logFC) obtenus lorsque l'on compare $\Delta bvgA$ /WT avec le mutant $bprC$ /WT (graphe de gauche) et lorsque l'on compare $\Delta bvgA$ /WT avec le double mutant $\Delta bvgA$ & mutant $bprC$ /WT (graphe de droite)

Une étude plus précise des gènes modulés montre que les *vag* non affectés par la diminution de BprC dans la cellule (cercle vert de la figure 7.4) sont des gènes classés comme des 'early' *vag* *i.e.* leur expression est activée rapidement après un switch de virulence (*bvg*- vers *bvg*+). On y trouve entre autre *fhaB*, *sphB1*, *cyaA*, *fim2*, *fimA*, *fimD*, *fimC*, *bipA*, *prn*, *cyaB*, *fhaS*, *bvgR*, *bvgS* et *bvgA*. A l'opposé les gènes présentant une modulation similaire entre le mutant $bprC$ et le $\Delta bvgA$ (cercle bleu de la figure 7.4, panneau de gauche) sont des 'late' *vag* *i.e.* leur expression est activée tardivement après un shift de virulence. On y trouve entre autre l'opéron *ptx-ptl*, *vag8*, *tcfA*, *brkA*, *bfrD*, *ompQ*, *hotAB*, les gènes du T3SS. Il semblerait donc que la diminution de BprC due à la mutation du promoteur de son gène entraîne un état de shift de virulence

intermédiaire correspondant à une concentration plus faible de BvgA phosphorylé dans la cellule (voir figure 1.4 en introduction).

Pour vérifier cette hypothèse, la concentration en BvgA-P dans la cellule lors d'un shift de *bvg-* à *bvg+* a été évaluée par Western blot en utilisant un anticorps anti BvgA Phospho-tag. Des cultures de cellules sur la souche sauvage et sur le mutant *bprC* à été maintenue à 20°C pendant plusieurs passages (3 semaines) puis ont été subitement placées à 37°C et suivies pendant plusieurs heures. La présence en BprC ainsi que la quantité de BvgA-P ont été monitorées. La figure 7.5 met en parallèle les quantités des différents acteurs.

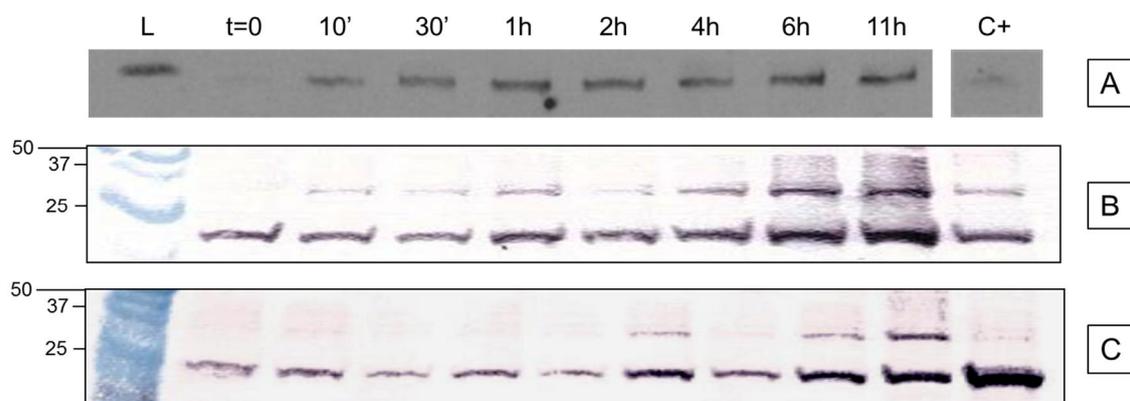


Figure 7.5 : Comparaison du profil d'expression de BprC et de BvgA-P lors d'un shift de phase *bvg-* à 20°C (t=0) à une phase *bvg+* à 37°C. Les temps correspondent aux temps de prélèvements post shift. (A) : Northern blot sur BprC dans la souche sauvage. (B) Western blot anti-BvgA-P (phospho-tag) dans la souche sauvage. (C) Western blot anti-BvgA-P (phospho-tag) dans la souche mutée *bprC*

La figure 7.5 montre une très nette augmentation de BvgA et de BvgA-P lors du shift de virulence. La bande attendue de 22,9 kDa ainsi que la bande supérieure correspondant à la forme phosphorylée augmente toutes deux de manière progressive pour atteindre un maximum 6h après le shift. Cette augmentation est nettement moins importante dans le mutant *bprC* ce qui semble indiquer que les profils de modulation des 'early' et 'late vag' observés en figure 7.4 sont dus à une plus faible concentration en BvgA et BvgA-P. La diminution de BprC dans la cellule entrainerait donc une incapacité pour la cellule d'accumuler assez de BvgA-P. BprC aurait donc un rôle à jouer sur le taux de traduction du gène *bvgA* soit en stabilisant l'ARNm de celui-ci ou en favorisant l'entrée des ribosomes sur le RBS. Ces hypothèses restent à valider.

CHAPITRE 8 : CONCLUSION - PERSPECTIVES

L'expertise que j'ai développée avec mon équipe au cours des années est centrée sur l'étude des caractéristiques génomiques au sens large d'un organisme. Cela va de l'étude de la composition et de l'organisation de son génome jusqu'à l'utilisation qui est faite de ce génome aussi bien pour des aspects de transcription-traduction que pour les aspects de régulation (*via* la dynamique du génome et/ou la production d'ARN régulateurs). Cette expertise a été mise à profit pour étudier la génomique de la bactérie pathogène *Bordetella pertussis* et essayer de mieux comprendre les relations entre ses propriétés génomiques et transcriptomiques et ses propriétés de virulence et d'adaptation.

Nos premières études, au début des années 2000, en utilisant la technologie des puces à ADN nous ont permis de mieux dimensionner le régulon principal de la virulence chez *Bordetella pertussis* en étudiant des conditions de modulation phénotypique et des mutants du système à deux composants BvgAS. Ces études ont montré à quel point les réseaux de régulation pouvaient être complexes aussi bien en terme de nombre de gènes régulés qu'en terme de diversité des fonctions de gènes impliqués. Nous avons déjà à l'époque évoqué le fait qu'un tel réseau ne pouvait être sous la régulation d'un seul et unique 'master' régulateur et que très probablement des sous régulons ou des régulateurs annexes devaient jouer leur rôle dans l'intégration de la réponse physiologique aux changements de conditions. Nous savons aujourd'hui que la régulation par le système BvgAS est interconnectée avec d'autres réseaux de régulation eux-mêmes sous le contrôle d'autres facteurs (Stenson et al, 2005 ; Croinin et al., 2005 ; Hoo et al., 2014). Il est à prévoir que la complexité des réseaux de régulation soit prochainement décortiquée par la mise en évidence de nouveaux facteurs et de leurs rôles. Ces facteurs seront probablement de nouveaux senseurs et/ou régulateurs protéiques mais aussi pour certains d'entre eux de petits ARN régulateurs (voir ci-dessous).

Une autre dimension de la régulation de la virulence a également été étudiée en utilisant les approches de puce à ADN au courant des années 2000. Cette régulation repose sur la dynamique des génomes et leur capacité à évoluer par diverses modification : perte ou gain de matériel génétique, remodelage de l'organisation du génome, petites variations (mutations, petites insertions, petites délétions). En comparant les contenus en gènes de plusieurs souches, isolées à différents moments et représentatives des grands phyla connus à l'époque, nous avons mis en évidence, en même temps que d'autres équipes, la force évolutive en action chez *Bordetella pertussis* qui repose sur l'élimination de parties de génome non essentielles au mode de vie qu'elle a choisi. Cette élimination progressive est la signature de la spécialisation de *pertussis* pour son hôte stricte, l'homme. D'autres aspects de la dynamique du génome

sont également très intéressants à étudier. Par exemple, la façon dont les parties de génome s'organisent entre elles et l'impact de cette organisation sur la virulence des souches a déjà été abordée (Brinig et al., 2006). Il est fort à parier que l'étude de l'impact des séquences d'insertion (IS), dont *Bordetella pertussis* est très riche, sur la régulation globale de la transcription et sur la régulation des gènes flanquants ces IS, serait de premier intérêt. Nos études pour établir le transcriptome primaire de la bactérie, dont une partie des résultats sont expliqués dans ce document, nous indiquent que la transcription de ces IS est particulière (résultats non montrés) et mériterait une attention plus importante.

En écho à la complexité des voies de régulation, notre étude du transcriptome primaire de *Bordetella pertussis* a révélé elle aussi un monde et une organisation d'une complexité insoupçonnée. La cartographie précise du transcriptome nous fait entrevoir une richesse de caractéristiques encore à peine étudiée, avec : (i) tout d'abord, un nombre de transcrits beaucoup plus important que ce que l'on pouvait imaginer grâce à la localisation de petits ARN entre les gènes annotés mais aussi en antisens des gènes codants, (ii) une cartographie de départ et de fin de gènes qui semble montrer des organisations parfois complexes comme des transcrits se chevauchants à leurs extrémités 5' ou 3' suggérant une interaction entre ces transcrits, (iii) des organisations en opérons avec des changements internes de niveaux de transcription, (iv) des structures particulières ayant très probablement une fonction de régulation telles que des riboswitches, des longs 5'UTR ou au contraire des ARNm sans leader.

Les niveaux de régulations 'classiques' de la cellule sont connus et documentés depuis de nombreuses années. On distingue la régulation transcriptionnelle, post-transcriptionnelle, traductionnelle et post-traductionnelle. La complexité du transcriptome telle qu'on commence à l'appréhender représente un nouveau niveau de régulation. Cette régulation est nettement moins documentée que les autres niveaux. Elle se fait au niveau de la synthèse, de l'utilisation et de la dégradation des ARN et elle passe par l'interaction d'ARN avec d'autre ARN tel que le fonctionnement des petits ARN régulateurs sur leurs cibles que sont, dans la plupart des cas, les ARN messagers.

L'équivalent eucaryote du petit ARN régulateur des bactéries est le 'small interfering RNA' (siRNA). Il se fixe sur un ARN messenger cible *via* le complexe RISC (RNA induce silencing complex) pour couper cette ARN messenger et ainsi provoquer sa dégradation par des exonucléases (Elbashir et al., 2001 ; Filipowicz et al., 2008). Cette propriété des siRNA leur a valu un engouement particulier lors de leur découverte car ils sont

apparus comme de potentielles nouvelles drogues très spécifiques pouvant cibler des produits de gènes habituellement difficiles à cibler. Les premières études ont cependant montré de forts effets 'hors-cible' qui ont refroidi les premiers espoirs (Robbins et al, 2008 ; Kleinman et al, 2008). Les difficultés d'utiliser de telles molécules en thérapie sont peu à peu levées notamment en ce qui concerne leur stabilité et leur distribution jusqu'aux cellules cibles (Peer et al, 2011). Il n'est pas irraisonnable de penser que des siRNA ou des dérivés de siRNA soient utilisés un jour en drogue thérapie. De même les petits ARN régulateurs bactériens pourraient être utilisés comme principe actif de futurs traitements antibiotiques. Ils ont pour eux leur très grande spécificité d'organismes étant donné qu'ils ne présentent qu'une très faible conservation de séquence même entre des souches bactériennes très proches.

Une autre voie d'étude des plus prometteuses sur ces petits ARN bactériens se situe au niveau de leur sécrétion par la bactérie afin de servir de signal vers d'autres bactéries ou d'effecteur vers les cellules de l'hôte infecté. Une étude sur *Listéria* (Abdullah et al. 2012) montre que la reconnaissance cytosolique de la bactérie passe par la sécrétion d'ARN et d'ADN bactérien par des bactéries vivantes dans le cytoplasme de la cellule infectée. Ces ARN et ADN sécrétés sont reconnus par des senseurs cytosoliques et déclenchent la production d'interféron β et l'activation de l'inflammation. Il est fort à parier que d'autres exemples de sécrétion vers les cellules de l'hôte impliquant des ARN codants mais aussi des non-codants vont voir le jour dans les prochains mois/années. Nous avons d'ores et déjà dans mon équipe réalisé, dans cette perspective de recherche, un RNA-seq du contenu des 'Outer Membranes Vesicles' (OMV) de *Bordetella pertussis* afin de mettre en évidence des ARN potentiellement véhiculés par la bactérie *via* ces vésicules.

Références

- Abdullah Z, Schlee M, Roth S, Mraheil MA, Barchet W, Bottcher J, Hain T, Geiger S, Hayakawa Y, Fritz JH, et al. 2012. RIG-I detects infection with live *Listeria* by sensing secreted bacterial nucleic acids. *EMBO J* 31: 4153–4164.
- Afonyushkin T, Vecerek B, Moll I, Bläsi U, Kaberdin VR. Both RNase E and RNase III control the stability of *sodB* mRNA upon translational inhibition by the small regulatory RNA RyhB. *Nucleic Acids Res.* 2005 Mar 21;33(5):1678-89.
- Akerley, B.J., Cotter, P.A., and Miller, J.F. (1995). Ectopic expression of the flagellar regulon alters development of the *Bordetella*-host interaction. *Cell* 80, 611–620.
- Akerley, B.J., Cotter, P.A., and Miller, J.F. (1995). Ectopic expression of the flagellar regulon alters development of the *Bordetella*-host interaction. *Cell* 80, 611–620.
- Alonso, S., Pethe, K., Mielcarek, N., Raze, D., and Locht, C. (2001). Role of ADP-ribosyltransferase activity of pertussis toxin in toxin-adhesin redundancy with filamentous hemagglutinin during *Bordetella pertussis* infection. *Infect. Immun.* 69, 6038–6043.
- Amman F, Wolfinger MT, Lorenz R, Hofacker IL, Stadler PF, Findeiß S. TSSAR: TSS annotation regime for dRNA-seq data. *BMC Bioinformatics.* 2014 Mar 27;15:89.
- Arico, B., Miller, J.F., Roy, C., Stibitz, S., Monack, D., Falkow, S., Gross, R., and Rappuoli, R. (1989). Sequences required for expression of *Bordetella pertussis* virulence factors share homology with prokaryotic signal transduction proteins. *Proc. Natl. Acad. Sci. USA* 86, 6671–6675.
- Arico, B., Nuti, S., Scarlato, V., and Rappuoli, R. (1993). Adhesion of *Bordetella pertussis* to eukaryotic cells requires a time-dependent export and maturation of filamentous hemagglutinin. *Proc. Natl. Acad. Sci. USA* 90, 9204–9208.
- Bartoloni, A., Pizza, M., Bigio, M., Nucci, D., Ashworth, L.A., Irons, L.I., Robinson, A., Burns, D., Manclark, C., Sato, H., and Rappuoli, R. (1988). Mapping of a protective epitope of pertussis toxin by in vitro refolding of recombinant fragments. *Biotechnology* 6, 709–712.
- Bibova I, Skopova K, Masin J, Cerny O, Hot D, Sebo P, Vecerek B. The RNA chaperone Hfq is required for virulence of *Bordetella pertussis*. *Infect Immun.* 2013 Nov;81(11):4081-90.
- Bjornstad, O.N., and Harvill, E.T. (2005). Evolution and emergence of *Bordetella* in humans. *Trends Microbiol.* 13, 355–359.
- Boucher, P.E., Murakami, K., Ishihama, A., and Stibitz, S. (1997). Nature of DNA binding and RNA polymerase interaction of the *Bordetella pertussis* BvgA transcriptional activator at the *fha* promoter. *J. Bacteriol.* 179, 1755–1763.
- Boucher, P.E., Yang, M.S., Schmidt, D.M., and Stibitz, S. (2001). Genetic and biochemical analyses of BvgA interaction with the secondary binding region of the *fha* promoter of *Bordetella pertussis*. *J. Bacteriol.* 183, 536–544.
- Brenner, D.J., Krieg, N.R., and Staley, J.T. (eds) (2005). Family III. Alcaligenaceae. In: *Bergey's Manual of Systematic Bacteriology, Part C.* Springer, pp. 647–671.
- Brinig MM, Cummings CA, Sanden GN, Stefanelli P, Lawrence A, Relman DA. Significant gene order and expression differences in *Bordetella pertussis* despite limited gene content variation. *J Bacteriol.* 2006 Apr;188(7):2375-82.
- Caldelari I, Chao Y, Romby P, Vogel J. RNA-mediated regulation in pathogenic bacteria. *Cold Spring Harb Perspect Med.* 2013 Sep 1;3(9):a010298.
- Caro V, Njamkepo E, Van Amersfoort SC, Mooi FR, Advani A, Hallander HO, He Q, Mertsola J, Riffelmann M, Vahrenholz C, Von König CH, Guiso N. Pulsed-field gel electrophoresis analysis of *Bordetella pertussis* populations in various European countries with different vaccine policies. *Microbes Infect.* 2005 Jun;7(7-8):976-82.
- Caro V., Elomaa A, Brun D, Mertsola J, He Q, Guiso N. (2006a). *Bordetella pertussis*, Finland and France. *Emerg. Infect. Dis.* 12, 987–989.
- Caro V., Hot D., Guigon G., Hubans C., Arrive M., Soubigou G., Renaud-Mongenie G., Antoine R., Locht C. Lemoine Y., Guiso N. (2006b). Temporal analysis of French *Bordetella pertussis* isolates by comparative whole-genome hybridization. *Microbes Infect.* In press.
- Chao Y, Vogel J. The role of Hfq in bacterial pathogens. *Curr Opin Microbiol.* 2010 Feb;13(1):24-33.

- Charles, I.G., Li, J.L., Roberts, M., Beesley, K., Romanos, M., Pickard, D.J., Francis, M., Campbell, D., Dougan, G., Brennan, M.J., Manclark, C.R., Jensen, M.A., Heron, I., Chubb, A., Novotny, P., and Fairweather, N.F. (1991). Identification and characterization of a protective immunodominant B cell epitope of pertactin (P.69) from *Bordetella pertussis*. *Eur. J. Immunol.* 21, 1147–1153.
- Cherry, J.D., Gornbein, J., Heininger, U., and Stehr, K. (1998). A search for serologic correlates of immunity to *Bordetella pertussis* cough illnesses. *Vaccine* 16, 1901–1906.
- Clantin, B., Hodak, H., Willery, E., Locht, C., Jacob-Dubuisson, F., and Villeret, V. (2004). The crystal structure of filamentous hemagglutinin secretion domain and its implications for the two-partner secretion pathway. *Proc. Natl. Acad. Sci. USA* 101, 6194–6199.
- Cotter PA, Jones AM. Phosphorelay control of virulence gene expression in *Bordetella*. *Trends Microbiol.* 2003 Aug;11(8):367-73.
- Cotter PA, Miller JF. Genetic analysis of the *Bordetella* infectious cycle. *Immunopharmacology.* 2000 Jul 25;48(3):253-5.
- Cotter, P.A., and DiRita, V.J. (2000). Bacterial virulence gene regulation: an evolutionary perspective. *Annu. Rev. Microbiol.* 54, 519–565.
- Coutte, L., Antoine, R., Drobecq, H., Locht, C., and Jacob-Dubuisson, F. (2001). Subtilisin-like autotransporter serves as maturation protease in a bacterial secretion pathway. *EMBO J.* 20, 5040–5048.
- Cróinín TO, Grippe VK, Merkel TJ. Activation of the *vrg6* promoter of *Bordetella pertussis* by *RisA*. *J Bacteriol.* 2005 Mar;187(5):1648-58.
- Croucher NJ, Thomson NR. Studying bacterial transcriptomes using RNA-seq. *Curr Opin Microbiol.* 2010 Oct;13(5):619-24.
- Cummings, C.A., Bootsma, H.J., Relman, D.A., and Miller, J.F. (2006). Species- and strain-specific control of a complex, flexible regulon by *Bordetella BvgAS*. *J. Bacteriol.* 188, 1775–1785.
- D'Argenio, D.A., and Miller, S.L. (2004). Cyclic di-GMP as a bacterial second messenger. *Microbiology* 150, 2497–2502.
- De Magistris, M.T., Di Tommaso, A., Domenighini, M., Censini, S., Tagliabue, A., Oksenberg, J.R., Steinman, L., Judd, A.K., O'Sullivan, D., and Rappuoli, R. (1992). Interaction of the pertussis toxin peptide containing residues 30–42 with DR1 and the T-cell receptors of 12 human T-cell clones. *Proc. Natl. Acad. Sci. USA* 89, 2990–2994.
- Diavatopoulos, D.A., Cummings, C.A., Schouls, L.M., Brinig, M.M., Relman, D.A., and Mooi, F.R. (2005). *Bordetella pertussis*, the causative agent of whooping cough, evolved from a distinct, human-associated lineage of *B. bronchiseptica*. *PLoS Pathog.* 1, e45.
- Dugar G, Herbig A, Förstner KU, Heidrich N, Reinhardt R, Nieselt K, Sharma CM. High-resolution transcriptome maps reveal strain-specific regulatory features of multiple *Campylobacter jejuni* isolates. *PLoS Genet.* 2013 May; 9(5):e1003495.
- Dupré E, Herrou J, Lensink MF, Wintjens R, Vagin A, Lebedev A, et al. (2015) Virulence Regulation with Venus Flytrap Domains: Structure and Function of the Periplasmic Moiety of the Sensor-Kinase BvgS. *PLoS Pathog* 11(3): e1004700.
- Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 2001; 411: 494–498.
- Elomaa, A., Advani, A., Donnelly, D., Antila, M., Mertsola, J., Hallander, H., and He, Q. (2005). Strain variation among *Bordetella pertussis* isolates in Finland, where the whole-cell pertussis vaccine has been used for 50 years. *J. Clin. Microbiol.* 43, 3681–3687.
- Felden B, Massire C, Westhof E, Atkins JF, Gesteland RF. Phylogenetic analysis of tmRNA genes within a bacterial subgroup reveals a specific structural signature. *Nucleic Acids Res.* 2001 Apr 1;29(7):1602-7.
- Filipowicz W, Bhattacharyya SN and Sonenberg N (2008). "Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?". *Nature Reviews Genetics* 9 (2): 102–114.
- Geissmann TA, Touati D. Hfq, a new chaperoning role: binding to messenger RNA determines access for small RNA regulator. *EMBO J.* 2004 Jan 28;23(2):396-405.
- Geuijen, C.A., Willems, R.J., Bongaerts, M., Top, J., Gielen, H., and Mooi, F.R. (1997). Role of the *Bordetella pertussis* minor fimbrial subunit, FimD, in colonization of the mouse respiratory tract. *Infect. Immun.* 65, 4222–4228. Boschwitz, J.S., van der Heide, H.G., Mooi, F.R., and Relman, D.A. (1997b). *Bordetella bronchiseptica* expresses the fimbrial structural subunit gene *fimA*. *J. Bacteriol.* 179, 7882–7885.
- Hannah, J.H., Menozzi, F.D., Renaud, G., Locht, C., and Brennan, M.J. (1994). Sulfated glycoconjugate receptors for the *Bordetella pertussis* adhesin filamentous hemagglutinin (FHA) and mapping of the heparin-binding domain on FHA. *Infect. Immun.* 62, 5010–5019.

- Hausman, S.Z., and Burns, D.L. (2000). Use of pertussis toxin encoded by *ptx* genes from *Bordetella bronchiseptica* to model the effects of antigenic drift of pertussis toxin on antibody neutralization. *Infect. Immun.* 68, 3763–3767.
- Hazenbos, W.L., van den Berg, B.M., and van Furth, R. (1993). Very late antigen-5 and complement receptor type 3 cooperatively mediate the interaction between *Bordetella pertussis* and human monocytes. *J. Immunol.* 151, 6274–6282.
- Heikkinen E, Kallonen T, Saarinen L, Sara R, King AJ, Mooi FR, Soini JT, Mertsola J, He Q. Comparative genomics of *Bordetella pertussis* reveals progressive gene loss in Finnish strains. *PLoS One.* 2007 Sep 19;2(9):e904.
- Hellwig, S.M., Hazenbos, W.L., van de Winkel, J.G., and Mooi, F.R. (1999). Evidence for an intracellular niche for *Bordetella pertussis* in broncho-alveolar lavage cells of mice. *FEMS Immunol. Med. Microbiol.* 26, 203–207.
- Hijnen, M., Mooi, F.R., van Gageldonk, P.G., Hoogerhout, P., King, A.J., and Berbers, G.A. (2004). Epitope structure of the *Bordetella pertussis* protein P.69 pertactin, a major Vaccine component and protective antigen. *Infect. Immun.* 72, 3716–3723.
- Hoo R, Lam JH, Huot L, Pant A, Li R, Hot D, Alonso S. Evidence for a role of the polysaccharide capsule transport proteins in pertussis pathogenesis. *PLoS One.* 2014 Dec 12;9(12):e115243.
- Hot D*, Bibova I*, Keidel K, Amman F, Slupek S, Cerny O, Gross R, Vecerek B. Transcriptional profiling of *Bordetella pertussis* reveals requirement of RNA chaperone Hfq for Type III secretion system functionality. *RNA Biol.* 2015 *: co-first authors
- Hot D, Slupek S, Wulbrecht B, D'Hondt A, Hubans C, Antoine R, Locht C, Lemoine Y. Detection of small RNAs in *Bordetella pertussis* and identification of a novel repeated genetic element. *BMC Genomics.* 2011 Apr 27;12(1):207.
- Hot, D., Antoine, R., Renauld-Mongenie, G., Caro, V., Hennuy, B., Levillain, E., Huot, L., Wittmann, G., Poncet, D., Jacob-Dubuisson, F., Guyard, C., Rimlinger, F., Aujame, L., Godfroid, E., Guiso, N., Quentin-Millet, M. J., Lemoine, Y., and Locht, C. (2003). Differential modulation of *Bordetella pertussis* virulence genes as evidenced by DNA microarray analysis. *Mol Genet Genomics* 269(4), 475-86.
- Jacob-Dubuisson, F., Locht, C., and Antoine, R. (2001). Two-partner secretion in Gram-negative bacteria: a thrifty, specific pathway for large virulence proteins. *Mol. Microbiol.* 40, 306–313.
- Jones, A.M., Boucher, P.E., Williams, C.L., Stibitz, S., and Cotter, P.A. (2005). Role of BvgA phosphorylation and DNA binding affinity in control of Bvg-mediated phenotypic phase transition in *Bordetella pertussis*. *Mol. Microbiol.* 58, 700–713.
- Khelef, N., Bachelet, C.M., Vargaftig, B.B., and Guiso, N. (1994). Characterization of murine lung inflammation after infection with parental *Bordetella pertussis* and mutants deficient in adhesins or toxins. *Infect. Immun.* 62, 2893–2900.
- Kimura, A., Mountzouros, K.T., Relman, D.A., Falkow, S., and Cowell, J.L. (1990). *Bordetella pertussis* filamentous hemagglutinin: evaluation as a protective antigen and colonization factor in a mouse respiratory infection model. *Infect. Immun.* 58, 7–16.
- King, A.J., Berbers, G., van Oirschot, H.F., Hoogerhout, P., Knipping, K., and Mooi, F.R. (2001). Role of the polymorphic region 1 of the *Bordetella pertussis* protein pertactin in immunity. *Microbiology* 147, 2885–2895.
- Kinnear, S.M., Marques, R.R., and Carbonetti, N.H. (2001). Differential regulation of Bvg-activated virulence factors plays a role in *Bordetella pertussis* pathogenicity. *Infect. Immun.* 69, 1983–1993.
- Kleinman ME, Yamada K, Takeda A, Chandrasekaran V, Nozaki M, Baffi JZ et al. Sequence- and target-independent angiogenesis suppression by siRNA via TLR3. *Nature* 2008; 452: 591–597.
- Lacey, B.W. (1960). Antigenic modulation of *Bordetella pertussis*. *J. Hyg. Camb.* 58, 57–93.
- Langmead B, Salzberg S. Fast gapped-read alignment with Bowtie 2. *Nature Methods.* 2012, 9:357-359.
- Laoide BM, Ullmann A. Virulence dependent and independent regulation of the *Bordetella pertussis* *cya* operon. *EMBO J.* 1990 Apr;9(4):999-1005.
- Lapin, J.H. (1943). Whooping cough Springfield, Illinois, Charles C. Thomas.
- Lechner, M., Schmitt, K., Bauer, S., Hot, D., Hubans, C., Levillain, E., Locht, C., Lemoine, Y. and Gross, R. (2009). Genomic island excisions in *Bordetella pertussis*. *BMC Microbiology*, 9:141.

- Link TM, Valentin-Hansen P, Brennan RG. Structure of *Escherichia coli* Hfq bound to polyribadenylate RNA. *Proc Natl Acad Sci U S A*. 2009 Nov 17;106(46):19292-7.
- Livny J. Bioinformatic discovery of bacterial regulatory RNAs using SIPHT. *Methods Mol Biol*. 2012;905:3-14.
- McLafferty M. A., Harcus D. R., Hewlett E. L. (1988) Nucleotide sequence and characterization of a repetitive DNA element from the genome of *Bordetella pertussis* with characteristics of an insertion sequence. *J. Gen. Microbiol.* 134:2297–2306.
- Magoc T, Wood D, Salzberg SL. EDGE-pro: Estimated Degree of Gene Expression in Prokaryotic Genomes. *Evolutionary Bioinformatics* vol.9, pp.127-136, 2013.
- Mandal M, Lee M, Barrick JE, Weinberg Z, Emilsson GM, Ruzzo WL, Breaker RR. A glycine-dependent riboswitch that uses cooperative binding to control gene expression. *Science*. 2004 Oct 8;306(5694):275-9. Erratum in: *Science*. 2004 Nov 26;306(5701):1477.
- Mann PB, Kennett MJ, Harvill ET. Toll-like receptor 4 is critical to innate host defense in a murine model of bordetellosis. *J Infect Dis*. 2004 Mar 1;189(5):833-6.
- Mann PB, Wolfe D, Latz E, Golenbock D, Preston A, Harvill ET. Comparative toll-like receptor 4-mediated innate host defense to *Bordetella* infection. *Infect Immun*. 2005 Dec;73(12):8144-52.
- Mao F, Dam P, Chou J, Olman V, Xu Y, DOOR: a Database of prokaryotic Operons, *Nucl. Acids Res*. 37: D459-D463, 2009.
- Massé E, Gottesman S. A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *Proc Natl Acad Sci U S A*. 2002 Apr 2;99(7):4620-5.
- Mattoo S, Foreman-Wykert AK, Cotter PA, Miller JF. Mechanisms of *Bordetella* pathogenesis. *Front Biosci*. 2001 Nov 1;6:E168-86.
- Mattoo, S., Miller, J.F., and Cotter, P.A. (2000). Role of *Bordetella bronchiseptica* fimbriae in tracheal colonization and development of a humoral immune response. *Infect. Immun.* 68, 2024–2033.
- Melvin JA, Scheller EV, Miller JF, Cotter PA. *Bordetella pertussis* pathogenesis: current and future challenges. *Nat Rev Microbiol*. 2014 Apr;12(4):274-88.
- Melvin JA, Scheller EV, Noël CR, Cotter PA. New Insight into Filamentous Hemagglutinin Secretion Reveals a Role for Full-Length FhaB in *Bordetella* Virulence. *MBio*. 2015 Aug 18;6(4).
- Mikulecky PJ, Kaw MK, Brescia CC, Takach JC, Sledjeski DD, Feig AL. *Escherichia coli* Hfq has distinct interaction surfaces for DsrA, rpoS and poly(A) RNAs. *Nat Struct Mol Biol*. 2004 Dec;11(12):1206-14.
- Moll I, Leitsch D, Steinhauser T, Bläsi U. RNA chaperone activity of the Sm-like Hfq protein. *EMBO Rep*. 2003 Mar;4(3):284-9.
- Mooi FR, van Loo IH, van Gent M, He Q, Bart MJ, Heuvelman KJ, de Greeff SC, Diavatopoulos D, Teunis P, Nagelkerke N, Mertsola J. *Bordetella pertussis* strains with increased toxin production associated with pertussis resurgence. *Emerg Infect Dis*. 2009 Aug;15(8):1206-13.
- Mooi, F.R., Jansen, W.H., Brunings, H., Gielen, H., van der Heide, H.G., Walvoort, H.C., and Guinee, P.A. (1992). Construction and analysis of *Bordetella pertussis* mutants defective in the production of fimbriae. *Microb. Pathog.* 12, 127–135.
- Mooi, F.R., van Loo, I.H., and King, A.J. (2001). Adaptation of *Bordetella pertussis* to vaccination, a cause for its reemergence? *Emerg. Infect. Dis.* 7, 526–528.
- Mooi, F.R., van Oirschot, H., Heuvelman, K., van der Heide, H.G., Gaastra, W., and Willems, R.J. (1998). Polymorphism in the *Bordetella pertussis* virulence factors P.69/pertactin and pertussis toxin in The Netherlands, temporal trends and evidence for vaccine-driven evolution. *Infect. Immun.* 66, 670–675.
- Musser, J.M., Bemis, D.A., Ishikawa, H., and Selander, R.K. (1987). Clonal diversity and host distribution in *Bordetella bronchiseptica*. *J. Bacteriol.* 169, 2793–2803.
- Musser, J.M., Hewlett, E.L., Pepler, M.S., and Selander, R.K. (1986). Genetic diversity and relationships in populations of *Bordetella* spp. *J. Bacteriol.* 166, 230–237.
- Ott A, Idali A, Marchais A, Gautheret D. NAPP: the Nucleic Acid Phylogenetic Profile Database. *Nucleic Acids Res*. 2012 Jan;40(Database issue)

- Packard, E.R., Parton, R., Coote, J.G., and Fry, N.K. (2004). Sequence variation and conservation in virulence-related genes of *Bordetella pertussis* isolates from the UK. *J. Med. Microbiol.* 53, 355–365.
- Parkhill J, Dougan G, James KD, Thomson NR, Pickard D, Wain J, Churcher C, Mungall KL, Bentley SD, Holden MT, Sebahia M, Baker S, Basham D, Brooks K, Chillingworth T, Connor P, Cronin A, Davis P, Davies RM, Dowd L, White N, Farrar J, Feltwell T, Hamlin N, Haque A, Hien TT, Holroyd S, Jagels K, Krogh A, Larsen TS, Leather S, Moule S, O'Gaora P, Parry C, Quail M, Rutherford K, Simmonds M, Skelton J, Stevens K, Whitehead S, Barrell BG. Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. *Nature*. 2001 Oct 25;413(6858):848-52.
- Parkhill, J., Sebahia, M., Preston, A., Murphy, L.D., Thomson, N., Harris, D.E., Holden, M.T., Churcher, C.M., Bentley, S.D., Mungall, K.L., Cerdeno-Tarraga, A.M., Temple, L., James, K., Harris, B., Quail, M.A., Achtman, M., Atkin, R., Baker, S., Basham, D., Bason, N., Cherevach, I., Chillingworth, T., Collins, M., Cronin, A., Davis, P., Doggett, J., Feltwell, T., Goble, A., Hamlin, N., Hauser, H., Holroyd, S., Jagels, K., Leather, S., Moule, S., Norberczak, H., O'Neil, S., Ormond, D., Price, C., Rabinowitsch, E., Rutter, S., Sanders, M., Saunders, D., Seeger, K., Sharp, S., Simmonds, M., Skelton, J., Squares, R., Squares, S., Stevens, K., Unwin, L., Whitehead, S., Barrell, B.G., and Maskell, D.J. (2003). Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*. *Nat. Genet.* 35, 32–40.
- Pedroni, P., Riboli, B., de Ferra, F., Grandi, G., Toma, S., Arico, B., and Rappuoli, R. (1988). Cloning of a novel pilin-like gene from *Bordetella pertussis*: homology to the fim2 gene. *Mol. Microbiol.* 2, 539–543.
- Peer D, Lieberman J. Special delivery: targeted therapy with small RNAs. *Gene Ther.* 2011 Dec;18(12):1127-33.
- Porwollik S, Boyd EF, Choy C, Cheng P, Florea L, Proctor E, McClelland M. Characterization of *Salmonella enterica* subspecies I genovars by use of microarrays. *J Bacteriol.* 2004 Sep;186(17):5883-98.
- Riboli, B., Pedroni, P., Cuzzoni, A., Grandi, G., and de Ferra, F. (1991). Expression of *Bordetella pertussis* fimbrial (fim) genes in *Bordetella bronchiseptica*: fimX is expressed at a low level and vir-regulated. *Microb. Pathog.* 10, 393–403.
- Robbins M, Judge A, Ambegia E, Choi C, Yaworski E, Palmer L et al. Misinterpreting the therapeutic effects of siRNA caused by immune stimulation. *Hum Gene Ther* 2008; 19: 991–999.
- Roy, C.R., and Falkow, S. (1991). Identification of *Bordetella pertussis* regulatory sequences required for transcriptional activation of the fhaB gene and autoregulation of the bvgAS operon. *J. Bacteriol.* 173, 2385–2392.
- Saukkonen, K., Cabellos, C., Burroughs, M., Prasad, S., and Tuomanen, E. (1991). Integrin-mediated localization of *Bordetella pertussis* within macrophages: role in pulmonary colonization. *J. Exp. Med.* 173, 1143–1149.
- Scarlato, V., Arico, B., Prugnola, A., and Rappuoli, R. (1991). Sequential activation and environmental regulation of virulence genes in *Bordetella pertussis*. *EMBO J.* 10, 3971–3975.
- Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, Chabas S, Reiche K, Hackermüller J, Reinhardt R, Stadler PF, Vogel J. The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature*. 2010 Mar 11;464(7286):250-5.
- Sledjeski DD, Whitman C, Zhang A. Hfq is necessary for regulation by the untranslated RNA DsrA. *J Bacteriol.* 2001 Mar;183(6):1997-2005.
- Sonnleitner E, Abdou L, Haas D. Small RNA as global regulator of carbon catabolite repression in *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A.* 2009 Dec 22;106(51):21866-71.
- Stenson TH, Allen AG, Al-Meer JA, Maskell D, Pepler MS. *Bordetella pertussis* risA, but not risS, is required for maximal expression of Bvg-repressed genes. *Infect Immun.* 2005 Sep;73(9):5995-6004.
- Stibitz, S., Aaronson, W., Monack, D., and Falkow, S. (1989). Phase variation in *Bordetella pertussis* by frameshift mutation in a gene for a novel two-component system. *Nature* 338, 266–269.
- Stibitz, S., and Yang, M.-S. (1991). Subcellular localization and immunological detection of proteins encoded by the vir locus of *Bordetella pertussis*. *J. Bacteriol.* 173, 4288–4293.
- Stibitz S. Mutations affecting the alpha subunit of *Bordetella pertussis* RNA polymerase suppress growth inhibition conferred by short C terminal deletions of the response regulator BvgA. *J Bacteriol.* 1998a May;180(9):2484-92.
- Stibitz S. IS481 and IS1002 of *Bordetella pertussis* create a 6-base-pair duplication upon insertion at a consensus target site. *J Bacteriol.* 1998b Sep;180(18):4963-6.

- Storsaeter, J., Hallander, H.O., Gustafsson, L., and Olin, P. (1998). Levels of anti-pertussis antibodies related to protection after household exposure to *Bordetella pertussis*. *Vaccine* 16, 1907–1916.
- Tsang, R.S., Lau, A.K., Sill, M.L., Halperin, S.A., Van Caesele, P., Jamieson, F., and Martin, I.E. (2004). Polymorphisms of the fimbria fim3 gene of *Bordetella pertussis* strains isolated in Canada. *J. Clin. Microbiol.* 42, 5364–5367.
- Tsui HC, Leung HC, Winkler ME. Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in *Escherichia coli* K-12. *Mol Microbiol.* 1994 Jul;13(1):35-49.
- Udekwi KI, Darfeuille F, Vogel J, Reimegård J, Holmqvist E, Wagner EG. Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA. *Genes Dev.* 2005 Oct 1;19(19):2355-66.
- Uhl, M.A., and Miller, J.F. (1994). Autophosphorylation and phosphotransfer in the *Bordetella pertussis* BvgAS signal transduction cascade. *Proc. Natl. Acad. Sci. USA* 91, 1163–1167.
- Uhl, M.A., and Miller, J.F. (1996a). Integration of multiple domains in a two-component sensor protein: the *Bordetella pertussis* BvgAS phosphorelay. *EMBO J.* 15, 1028–1036.
- Uhl, M.A., and Miller, J.F. (1996b). Central role of the BvgS receiver as a phosphorylated intermediate in a complex two-component phosphorelay. *J. Biol. Chem.* 271, 33176–33180.
- Urisu, A., Cowell, J.L., and Manclark, C.R. (1986). Filamentous hemagglutinin has a major role in mediating adherence of *Bordetella pertussis* to human WiDr cells. *Infect. Immun.* 52, 695–701.
- van den Berg, B.M., Beekhuizen, H., Mooi, F.R., and van Furth, R. (1999a). Role of antibodies against *Bordetella pertussis* virulence factors in adherence of *Bordetella pertussis* and *Bordetella parapertussis* to human bronchial epithelial cells. *Infect. Immun.* 67, 1050–1055.
- van den Berg, B.M., Beekhuizen, H., Willems, R.J., Mooi, F.R., and van Furth, R. (1999b). Role of *Bordetella pertussis* virulence factors in adherence to epithelial cell lines derived from the human respiratory tract. *Infect. Immun.* 67, 1056–1062.
- Van der Zee, A., Mooi, F., Van Embden, J., and Musser, J. (1997). Molecular evolution and host adaptation of *Bordetella* spp.: phylogenetic analysis using multilocus enzyme electrophoresis and typing with three insertion sequences. *J. Bacteriol.* 179, 6609–6617.
- Van Loo, I.H., and Mooi, F.R. (2002). Changes in the Dutch *Bordetella pertussis* population in the first 20 years after the introduction of whole-cell vaccines. *Microbiology* 148, 2011–2018.
- Van Loo, I.H., Heuvelman, K.J., King, A.J., and Mooi, F.R. (2002). Multilocus sequence typing of *Bordetella pertussis* based on surface protein genes. *J. Clin. Microbiol.* 40, 1994–2001.
- Van Loo, I.H., van der Heide, H.G., Nagelkerke, N.J., Verhoef, J., and Mooi, F.R. (1999). Temporal trends in the population structure of *Bordetella pertussis* during 1949–1996 in a highly vaccinated population. *J. Infect. Dis.* 179, 915–923.
- Vandebriel, R.J., Hellwig, S.M., Vermeulen, J.P., Hoekman, J.H., Dormans, J.A., Roholl, P.J., and Mooi, F.R. (2003). Association of *Bordetella pertussis* with host immune cells in the mouse lung. *Microb. Pathog.* 35, 9–29.
- van Vliet AH. Next generation sequencing of microbial transcriptomes: challenges and opportunities. *FEMS Microbiol Lett.* 2010 Jan;302(1):1-7.
- Veal-Carr, W.L., and Stibitz, S. (2005). Demonstration of differential virulence gene promoter activation in vivo in *Bordetella pertussis* using RIVET. *Mol. Microbiol.* 55, 788–798.
- Vecerek B, Moll I, Bläsi U. Control of Fur synthesis by the non-coding RNA RyhB and iron-responsive decoding. *EMBO J.* 2007 Feb 21;26(4):965-75.
- Weber, C., Boursaux-Eude, C., Coralie, G., Caro, V., and Guiso, N. (2001). Polymorphism of *Bordetella pertussis* isolates circulating for the last 10 years in France, where a single effective whole-cell vaccine has been used for more than 30 years. *J. Clin. Microbiol.* 39, 4396–4403.
- Wek RC, Hatfield GW. Nucleotide sequence and in vivo expression of the ilvY and ilvC genes in *Escherichia coli* K12. Transcription from divergent overlapping promoters. *J Biol Chem.* 1986 Feb 15;261(5):2441-50.
- Willems, R.J., van der Heide, H.G., and Mooi, F.R. (1992). Characterization of a *Bordetella pertussis* fimbrial gene cluster which is located directly downstream of the filamentous haemagglutinin gene. *Mol. Microbiol.* 6, 2661–2671.

Winkler WC, Cohen-Chalamish S, Breaker RR. An mRNA structure that controls gene expression by binding FMN. *Proc Natl Acad Sci U S A*. 2002 Dec 10;99(25):15908-13. Epub 2002 Nov 27.

Yuk, M.H., Harvill, E.T., and Miller, J.F. (1998). The BvgAS virulence control system regulates type III secretion in *Bordetella bronchiseptica*. *Mol. Microbiol.* 28, 945–959.

Zhang A, Altuvia S, Tiwari A, Argaman L, Hengge-Aronis R, Storz G. The OxyS regulatory RNA represses rpoS translation and binds the Hfq (HF-I) protein. *EMBO J*. 1998 Oct 15;17(20):6061-8.