N°: 2016 AGPT 0007

Doctorat AgroParisTech

T H È S E

pour obtenir le grade de docteur délivré par

L'Institut des Sciences et Industries du Vivant et de l'Environnement

(AgroParisTech)

Spécialité : microbiologie

présentée et soutenue publiquement par

Sébastien GÉLIS-JEANVOINE

le 25 janvier 2016

Étude de la kurstakine et de Spo0A, deux facteurs d'adaptation de *Bacillus thuringiensis*

Directeur de thèse : **Didier LERECLUS**
Co-encadrement de la thèse : **Leyla SLAMTI**

Jury

Mme Véronique BROUSSOLLE, CR1, UMR 408 SQPOV, INRA PACA	Rapporteuse
M. Thomas CANDELA, MC, EA 4043 EMDS, Université Paris-Sud XI	Rapporteur
M. Alain GIVAUDAN, DR2, UMR 1333 DGIMI, INRA Montpellier	Examinateur
M. Philippe JACQUES, PR, Institut Charles Viollette, Polytech-Lille & Université Lille I	Examinateur
Mme Isabelle MARTIN-VERSTRAETE, PR, Laboratoire PBA, Institut Pasteur Paris	Examinatrice
M. Didier LERECLUS, DREx, UMR 1319 MICALIS, INRA Jouy-en-Josas	Directeur
Mme Leyla SLAMTI, CR1, UMR 1319 MICALIS, INRA Jouy-en-Josas	Co-encadrante

AgroParisTech
UMR 1319 MICALIS
Domaine de Vilvert – 78350 Jouy-en-Josas

« Moi je n'étais rien, et voilà qu'aujourd'hui
Je suis le gardien du sommeil de ses nuits,
Je l'aime à mourir. »

À Anne-Sophie et Zoé.

RÉSUMÉ / ABSTRACT

Bacillus thuringiensis est une bactérie entomopathogène du groupe *Bacillus cereus*. En plus de sa capacité à infecter et tuer les larves d'insectes sensibles, *B. thuringiensis* peut survivre à l'état végétatif plus d'une semaine dans le cadavre de l'hôte qu'il vient de tuer ; un mode de développement appelé nécrotrophisme. Comme d'autres *Bacillus*, *B. thuringiensis* possède aussi plusieurs mécanismes d'adaptation : formation de biofilm, sporulation et synthèse de métabolites secondaires originaux, par exemple. Ceux-ci sont finement régulés en fonction de l'environnement, *via* des systèmes d'intégration de signaux extra-cellulaires qui convergent souvent vers des régulateurs transcriptionnels centraux, parmi lesquels Spo0A et NprR. Spo0A est notamment présent chez les espèces *Bacillus* et *Clostridium*, et contrôle un grand nombre de mécanismes d'adaptation. NprR est un régulateur de *quorum sensing* présent uniquement dans le groupe *Bacillus cereus*. Chez *B. thuringiensis*, lorsque le *quorum* est atteint, il permet la transition d'un mode de vie pathogène à nécrotrophe. Le travail mené au cours de cette thèse a eu pour but de mieux comprendre le comportement de *B. thuringiensis* en phase stationnaire et au cours de l'infection d'un insecte hôte.

NprR régule la transcription d'une quarantaine de gènes, parmi lesquels le locus *krs*. Ce grand locus de 30 kb porte six gènes impliqués dans la synthèse non-ribosomique de la kurstakine, un lipo-heptapeptide susceptible de jouer un ou des rôle(s) importants dans le nécrotrophisme et revêtant un potentiel intérêt bio-technologique. Nous avons montré que les gènes sont co-transcrits à partir d'un promoteur en amont du premier gène, formant donc un opéron. La mesure de l'activité de ce promoteur nous a permis de mettre en évidence que celle-ci est fortement dépendante des conditions de croissance, et que sa régulation est complexe et fait intervenir, en plus de NprR, Spo0A et une séquence d'ADN palindromique dont la fonction est encore inconnue. Nous démontrons aussi que le produit de ce locus, la kurstakine, est impliqué dans la structuration du biofilm de *B. thuringiensis*.

Dans le cadre d'une approche plus globale, nous avons caractérisé les transcriptomes de *B. thuringiensis in vitro*, à deux points dans le temps, et *in vivo*, dans la souche modèle 407 et le mutant de délétion de *spo0A*, en utilisant la technique de RNA-seq. Couplées à une prédiction *in silico* des boîtes Spo0A dans le génome de *B. thuringiensis*, ces données décrivent pour la première fois le régulon Spo0A d'une bactérie du groupe *Bacillus cereus*. Nous tentons également d'élucider le rôle de Spo0A dans le cycle infectieux de *B. thuringiensis*, et montrons que sans *spo0A*, *B. thuringiensis* est incapable de persister dans le cadavre de son hôte. Ce défaut de survie est indépendant de la sporulation, car un mutant affecté dans les étapes plus tardives de la sporulation survit. Cela suggère qu'une partie du régulon Spo0A est également impliquée dans mode de vie nécrotrophe de *B. thuringiensis*.

Mots clés : *Bacillus thuringiensis*, Régulation transcriptionnelle, Nécrotrophisme, Spo0A, Kurstakine, Biofilm

Bacillus thuringiensis is an entomopathogenic bacterium from the *Bacillus cereus* group. Additionally to its ability to kill larvae of susceptible insects, *B. thuringiensis* is able to maintain a vegetative population in its dead host for more than a week; a behavior called necrotrophism. Like other *Bacillus*, *B. thuringiensis* is also capable of triggering several adaptation processes: biofilm formation, sporulation and synthesis of original secondary metabolites for instance. These processes are tightly regulated by environmental cues sensed by several sensors of extra-cellular signals, often linked to master regulators of transcription, among which are Spo0A and NprR. Spo0A is notably present in the *Bacillus* and *Clostridium* species and controls numerous adaptation processes. NprR is a *quorum sensing* regulator specific to the *Bacillus cereus* group. In *B. thuringiensis*, when the *quorum* is reached, it allows the transition between a pathogenic and a necrotrophic lifestyle. The work undertaken during this thesis aimed at a better understanding of the behavior of *B. thuringiensis* during the stationary phase and the infection of a host insect.

NprR regulates the transcription of about forty genes, among which are those of the *krs* locus. This big 30 kb locus harbors six genes involved in the non-ribosomal synthesis of kurstakin, a lipo-heptapeptide possibly playing an important role in necrotrophism, and exhibiting biotechnological potential. We showed that the genes are co-transcribed from a promoter upstream of the first gene, forming an operon. Measures of this promoter's activity revealed a strong dependence to growth conditions and a complex regulation involving, in addition to NprR, Spo0A and a DNA palindrome of yet unknown function. We also show that the product of this locus, kurstakin, is involved in the structuring of the biofilm of *B. thuringiensis*.

Using a broader approach, we characterized the *in vitro* (two time points) and *in vivo* transcriptomes of our model strain 407 and its *spo0A*-deletion mutant, using RNA-seq. Coupled to an *in silico* prediction of the Spo0A boxes in the genome of *B. thuringiensis*, these data describe for the first time the Spo0A regulon of a member of the *Bacillus cereus* group. We also investigate the role of Spo0A in the infection cycle of *B. thuringiensis* and show that a *spo0A*-less mutant of *B. thuringiensis* is unable to persist in its host cadaver. This survival defect does not depend on sporulation, because a mutant impaired in later steps of sporulation still has the ability to survive. This suggests that part of the Spo0A regulon is also involved in the necrotrophic behavior of *B. thuringiensis*.

Keywords: *Bacillus thuringiensis*, Transcriptional regulation, Necrotrophism, Spo0A, Kurstakin, Biofilm

REMERCIEMENTS

Je tiens en premier lieu à remercier Didier Lereclus pour la direction de cette thèse, et Leyla Slamti pour son encadrement. Je remercie également Véronique Broussolle, Thomas Candela, Alain Givaudan, Philippe Jacques et Isabelle Martin-Verstraete d'avoir accepté d'évaluer ce travail. Merci aux membres de mon comité de thèse : Arnaud Chastanet, Alain Givaudan et Jean-Marc Nicaud, nos réunions ont été très fructueuses. Merci au laboratoire ProBioGEM de Polytech Lille, et particulièrement à Thibault Caradec et Philippe Jacques, pour leur collaboration au projet BtSurf. Merci aussi à Christelle Lemy, dont je discute ici quelques résultats, pour sa collaboration sur ce même projet. Enfin, merci à l'école doctorale ABIES, spécialement Irina et Corinne, pour leur soutien tout au long de ces trois années

Je souhaite ensuite exprimer ma gratitude à tous ceux qui ont facilité mon séjour dans l'équipe GME, permanents, anciens et nouveaux.

J'ai aussi une pensée pour les nombreux collègues INRA, MICALIS, beaucoup mais pas que, avec qui j'ai pu avoir des discussions enrichissantes.

Enfin, un merci tout particulier à ma famille, celle d'en haut (Papa, Maman, Papys et Mamies), d'en bas (Zozo), du côté (Anne-So et Coco), et autres diagonales, pour leur soutien. Pour deux d'entre vous il est quotidien, et je ne vous dirai jamais assez à quel point j'en sors grandi. Papa, Maman, merci pour tout, vraiment tout. Coco… Bon courage ! C'est toi la prochaine :-)

ABRÉVIATIONS

A (domaine) : domaine d’adénylation.

AFLP : « amplified fragment length polymorphism ».

Ala : alanine.

AMPc : adénosine monophosphate cyclique.

API : « application programming interface ».

ARN : acide ribonucléique.

ARNm : ARN messager.

ARNr : ARN ribosomique.

ATCC : « American type culture collection ».

ATP : adénosine tri-phosphate.

B. cereus : *Bacillus cereus sensu stricto*.

BHI : « brain-heart infusion ».

BLAST : « basic local alignment search tool ».

B. thuringiensis : *Bacillus thuringiensis*.

C (domaine) : domaine de condensation.

CDS : « coding DNA sequence ».

ChIP-seq : « chromatin immuno-precipitation sequencing ».

CMC : concentration micellaire critique.

COG : « cluster of orthologous groups ».

D – (devant une molécule) : dextrogyre.

di-GMPc : di–GMP cyclique.

E (domaine) : domaine d’épimérisation.

EBI : « European bioinformatics institute ».

GC : guanine et cytosine.

GFP : « green fluorescent protein ».

Gln : glutamine.

Gly : glycine.

GMP : guanosine mono-phosphate.

GO : « gene ontology ».

GTK : « GIMP toolkit ».

HCT : hydrolysat de caséine tryptone.

HSP : « high-scoring segment pair »

His : histidine.

HTH (motif) : hélice-tour-hélice.

kb : kilo-base, soit 10^3 bases (nucléotides).

L – (devant une molécule) : lévogyre.

LB : « lysogeny broth ».

LBP : LB peptone.

Mb : méga-base, soit 10^6 bases (nucléotides).

MLEE : « multi-locus enzyme electrophoresis ».

MLST : « multi-locus sequence typing ».

NCBI : « national center for biotechnology information ».

NRPS : « non-ribosomal peptide synthetase ».

nt : nucléotide.

PCP (domaine) : domaine « peptidyl carrier protein ».

PDB : « protein data bank ».

PHB : poly(3-hydroxybutyrate).

POT : « proton-dependent oligopeptide transporter ».

PPant : 4'-phosphopantéthéine.

PPTase : 4'-phosphopantéthéinyl transférase.

PPi : ion pyrophosphate ($P_2O_7^{4-}$).

qPCR : « quantitative polymerase chain reaction »

REST (API) : « representational state transfer ».

RNA-seq : « RNA-sequencing ».

RNPP (famille) : Rap, NprR, PlcR, PrgX.

Ser : Sérine.

SHA : « secure hash algorithm ».

SNP : « single nucleotide polymorphism ».

SOAP (API) : « simple object access protocol ».

Te : thioestérase.

TIAC : toxi-infection alimentaire collective.

Thr : thréonine.

UT : unité transcriptionnelle (gène ou opéron).

WT : « wild type ».

Liste des figures

Figure 1 : Phylogénie ARNr 16S du groupe *Bacillus cereus*.

Figure 2 : Phylogénie MLST du groupe *Bacillus cereus*.

Figure 3 : Spore de *Bacillus thuringiensis* et son cristal.

Figure 4 : Schéma de l'infection d'une larve sensible par *Bacillus thuringiensis*.

Figure 5 : Exemples de différenciations bactériennes.

Figure 6 : différenciations au cours du cycle infectieux de *Bacillus thuringiensis* dans *Galleria mellonella*.

Figure 7 : Schéma général des systèmes de *quorum sensing* utilisant des peptides.

Figure 8 : Modèle du mécanisme de répression transcriptionnelle de l'opéron *tapA–sipW–tasA* par SinR.

Figure 9 : Modèle de l'interaction SinR–SlrR.

Figure 10 : Modèle d'activation de PlcR.

Figure 11 : Modèle du tétramère de NprR lié à l'ADN.

Figure 12 : Schéma général de l'activation de Spo0A.

Figure 13 : *spo0A* et son promoteur chez *Bacillus subtilis* 168.

Figure 14 : Les boucles de rétro-contrôle impliquées dans la transcription et l'activation de Spo0A.

Figure 15 : Schéma du modèle de phosphorelais chez *Bacillus subtilis*.

Figure 16 : Accumulation de Spo0A~P et différenciations de *Bacillus subtilis*.

Figure 17 : Modèle du cannibalisme chez *Bacillus subtilis*.

Figure 18 : Modèle de régulation de la formation de biofilm chez *Bacillus subtilis*.

Figure 19 : Représentants des 4 familles de lipopeptides de *Bacillus*.

Figure 20 : Structure annotée de kurstakine cyclique avec acide gras en C12 β-hydroxylé.

Figure 21 : Action séquentielle de la surfactine sur une bicouche lipidique mixte.

Figure 22 : Modèle de la signalisation par la surfactine.

Figure 23 : Les domaines A, PCP et C des modules NRPS.

Figure 24 : Le locus *krs* de *Bacillus thuringiensis* 407.

Figure 25 : Workflow de BLASTats.

Figure 26 : Interface de BLASTats.

Figure 27 : Diagramme circulaire des fonctions COG portées par le génome de *B. thuringiensis* 407.

Figure 28 : Interface de PyStrains.

Figure 29 : Spo0A régule la bimodalité d'expression de *sinI* dans la population.

SOMMAIRE

Résumé / Abstract....3
Remerciements....6
Abréviations....8
Liste des figures....12
Sommaire....15
Introduction bibliographique....19
Chapitre I - Bacillus cereus sensu lato....20
1 Généralités....20
2 Le cycle infectieux de *Bacillus thuringiensis*....24
2.1 Destruction de la barrière intestinale et invasion des tissus....24
2.2 Nécrotrophisme....25
2.3 Dissémination....26
Chapitre II - différenciations bactériennes....28
1 Les bactéries aussi se différencient....28
2 Les signaux déclenchant la différenciation....30
2.1 Les signaux « professionnels » : le *quorum sensing*....30
2.2 Les messagers secondaires....31
3 Les régulateurs de la différenciation....31
3.1 Spo0A : le régulateur en chef....31
3.2 Les facteurs sigma....32
3.3 CodY....32
3.4 AbrB....33
3.5 SinR....35
3.6 PlcR....37
3.7 NprR....37
Chapitre III - Spo0A, le régulateur clé de la différenciation....41
1 Schéma général....41
2 Régulation transcriptionnelle de *spo0A*....43
3 Régulation post-traductionnelle de Spo0A....45
3.1 Les kinases....45
3.2 Le relais et les phosphatases....46
4 Les voies de différenciation régulées par Spo0A~P....48
4.1 Une question de concentration....48
4.2 La mobilité par « sliding »....50
4.3 Le cannibalisme....50

4.4 La formation de biofilm....52
4.5 La sporulation....53
4.6 La production de toxines....54
5 Et chez *Bacillus thuringiensis* ?....55
Chapitre IV - La kurstakine....58
1 Les lipopeptides de *Bacillus*....58
2 Découverte de la kurstakine....62
3 La voie de synthèse non-ribosomique....62
4 Les NRPS de *Bacillus thuringiensis* 407....66
5 Le locus *krs*....66
Contexte de l'étude....68
Résultats....71
Chapitre I - Transcription et rôle de la kurstakine....72
1 Article 1 : « Transcription and role of kurstakin, a lipopeptide produced by *Bacillus thuringiensis* »....72
1.1 Phylogénie du locus *krs*....72
1.2 La régulation transcriptionnelle de l'opéron *krs* est complexe....73
1.3 La kurstakine est impliquée dans la structuration du biofilm....74
Article 1 : « Transcription and role of kurstakin, a lipopeptide produced by *Bacillus thuringiensis* »....75
Chapitre II - Le régulon Spo0A de *Bacillus thuringiensis* est nécessaire à son nécrotrophisme....76
1 Article 2 : « The Spo0A regulon of *Bacillus thuringiensis* is required for its necrotrophism »....76
1.1 *spo0A* est nécessaire au nécrotrophisme de *B. thuringiensis*....76
1.2 Le régulon Spo0A de *B. thuringiensis*....77
1.3 Le sous-régulon potentiellement impliqué dans le nécrotrophisme....78
Article 2 : « The Spo0A regulon of *Bacillus thuringiensis* is required for its necrotrophism »....80
Chapitre III - Logiciels et bases de données développés pendant la thèse....82
1 BLASTats....82
1.1 Généralités....82
1.2 Développement....82
1.3 Fonctionnement....84
2 EggnogGenome....85
2.1 Généralités....85
2.2 Développement....87

2.3 Fonctionnement....87
3 PyStrains....89
3.1 Généralités....89
3.2 Développement....89
3.3 Fonctionnement....89
4 Bases de données....90
Discussion et perspectives....92
Chapitre I - La kurstakine....93
1 La régulation de la production de kurstakine....93
1.1 Régulation transcriptionnelle de *krs*....93
1.1.a Le rôle exact du palindrome....93
1.1.b Les *trans*-régulateurs....93
1.2 Régulation post-transcriptionnelle....94
2 Le mode d'action de la kurstakine dans la structuration du biofilm....96
3 Les autres rôles de la kurstakine....97
3.1 Applications....98
Chapitre II - Spo0A chez *Bacillus thuringiensis*....100
1 Les sites réels de fixation de Spo0A sur l'ADN....100
2 Les régulations spécifiques à *B. thuringiensis*....100
2.1 La sporulation....100
2.2 La formation de biofilm....102
2.3 La mobilité....103
2.4 Les potentiels régulateurs transcriptionnels sous le contrôle de Spo0A....104
2.5 Les plasmides....104
3 Caractérisation des effecteurs supposés importants pour le nécrotrophisme....105
Valorisation de la thèse....106
Références bibliographiques....108
Annexes....126
Erratum sur l'article : « Necrotrophism is a quorum-sensing-regulated lifestyle in *Bacillus thuringiensis*. »....127

INTRODUCTION BIBLIOGRAPHIQUE

Chapitre I - Bacillus cereus sensu lato

1 Généralités

Dans le phylum des *Firmicutes*, le genre *Bacillus* contient un groupe de 8 espèces, le groupe *Bacillus cereus* (*sensu lato* pour le distinguer de l'espèce *Bacillus cereus sensu stricto*). Ce groupe contient actuellement les espèces *Bacillus anthracis*, *Bacillus cereus*, *Bacillus cytotoxicus*, *Bacillus mycoides*, *Bacillus pseudomycoides*, *Bacillus thuringiensis*, *Bacillus toyonensis* et *Bacillus weihenstephanensis*.

Les bactéries du groupe *Bacillus cereus* sont des bacilles à coloration de Gram positive, aérobies facultatives, mobiles ou immobiles, et capables de former des spores. Elles ont un génome pauvre en GC (autour de 35 %) composé d'un chromosome long de 4 à 5.5 Mb et souvent de plasmides de tailles très variables, entre 2 et 400 kb.

Bacillus anthracis est l'agent causal de la maladie du charbon (anthrax). Cette maladie, décrite dans la littérature depuis plus de 2000 ans, n'a été imputée formellement à *B. anthracis* qu'en 1876 par Robert Koch (Koch, 1876; Moïse, 700BC).

Bacillus cereus sensu stricto (*B. cereus*) est un pathogène humain responsable de toxi-infections alimentaires collectives (TIAC). En 2008, *B. cereus* aurait été la troisième cause de TIAC en France, même si sa symptomatologie proche d'autres pathogènes alimentaires laisse supposer que le nombre de cas est sous-estimé (ANSES, 2011).

L'espèce *Bacillus cytotoxicus* a été proposée en 2008, afin de séparer la souche NVH 391-98 de *B*c (Lapidus *et al.*, 2008). Cette souche diffère par plusieurs aspects des souches « classiques » de *B. cereus* : haute pathogénie, chromosome réduit (4 Mb *versus* 5,5 Mb), distance phylogénétique aux souches de *Bc* élevée, absence du facteur sigma σ^B notamment.

Bacillus mycoides a été décrite pour la première fois en 1886 (Flügge). Elle contenait en fait à l'origine deux espèces, qui ont ensuite été séparées en *Bacillus mycoides* et *Bacillus pseudomycoides* (Nakamura, 1998). Ces deux espèces se distinguent de *Bc* par leur absence de motilité et la morphologie filamenteuse de leurs colonies qui leur donne leur nom. *B. mycoides* se distingue de *B. pseudomycoides* par la composition en acide gras de sa membrane.

Fig. 1 : Phylogénie ARNr 16S du groupe *Bacillus cereus*.

L'arbre contient la séquence de 169 ARNr 16S du groupe *Bacillus cereus* (http://rdp.cme.msu.edu/) et a été construit par approximation du maximum de vraisemblance avec FastTree (Price *et al.*, 2010).

Rouge : *B. anthracis*,
bleu : *B. cereus*,
gris : *B. cytotoxicus*,
vert : *B. thuringiensis*,
cyan : *B. toyonensis*,
ocre : *B. mycoides*,
fushia : *B. pseudomycoides*,
orange : *B. weihenstephanensis*.

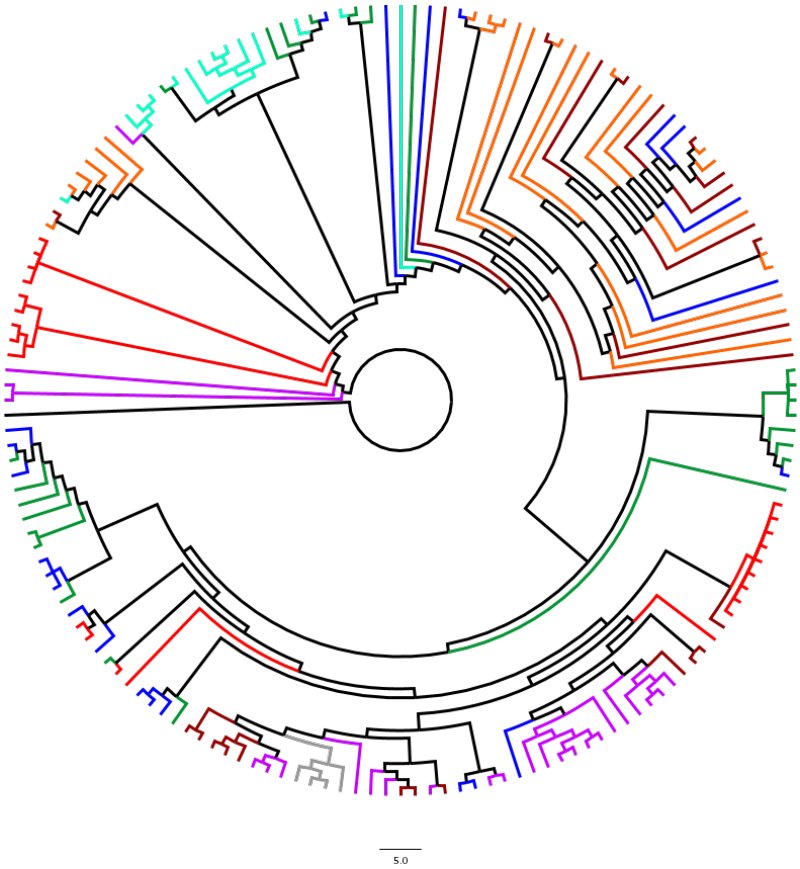
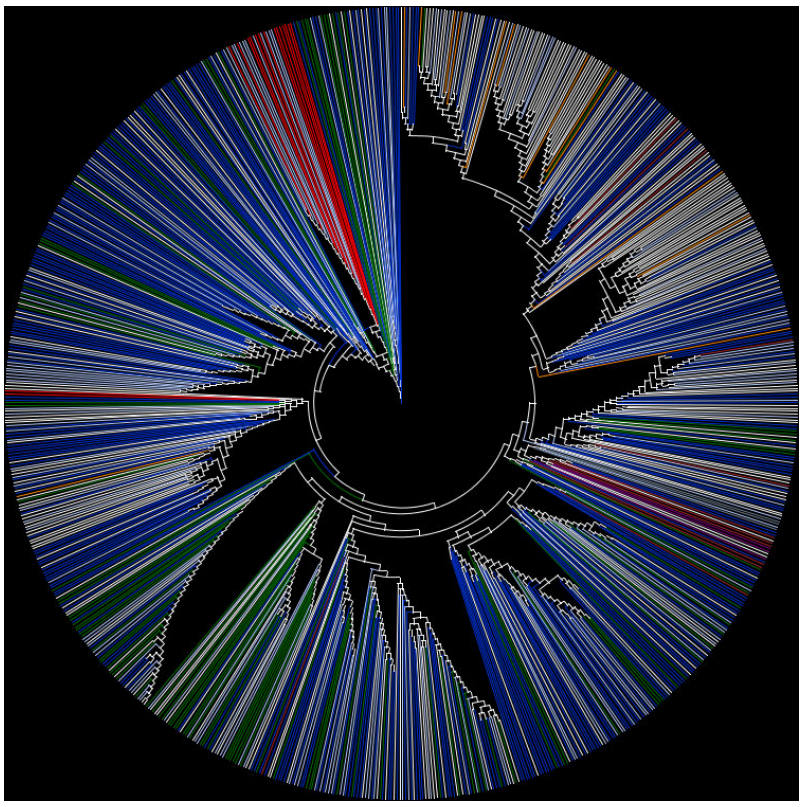


Fig. 2 : Phylogénie MLST du groupe *Bacillus cereus*.

L'arbre contient 1484 souches et est construit par maximum de vraisemblance à partir de la séquence de 26 gènes de ménage (http://mlstoslo.uio.no/) (Tourasse and Kolstø, 2008).

Rouge : *B. anthracis*,
bleu : *B. cereus*,
gris : *B. cytotoxicus*,
vert : *B. thuringiensis*,
ocre : *B. mycoides*,
fushia : *B. pseudomycoides*,
orange : *B. weihenstephanensis*.



Bacillus thuringiensis (*B. thuringiensis*) a été isolé parallèlement au Japon et en Allemagne par deux chercheurs, Shigetani Ishiwata et Ernst Berliner, entre 1901 et 1911. Il a rapidement été établi que *Bt* était entomopathogène (pathogène d'insecte), puis, plus tard, que cette propriété dépendait de cristaux toxiques qu'il produit (Angus, 1953, 1954; Aoki and Chigasaki, 1915; Berliner, 1911, 1915; Hannay, 1953). Sa toxine Cry1Ab est notamment utilisée dans le maïs transgénique MON810, et rend le plant résistant à la pyrale (U.S. Environmental Protection Agency, 2010).

L'espèce *Bacillus toyonensis* a été proposée en 2013 et contient une souche ayant appartenu précédemment à *B. cereus* (Jiménez *et al.*). Deux critères ont été retenus pour avancer la création d'une espèce distincte : distance génomique et protéomique aux autres souches de *Bc*, et propriété probiotique de la souche.

Bacillus weihenstephanensis est une autre espèce issue d'une scission au sein de *Bc* (Lechner *et al.*, 1998). La caractéristique discriminante de cette espèce était initialement la psychrotolérance, ces souches étant capables de se développer en deçà de 7 °C. Cependant, des souches de *B. cereus* psychrotolérantes ont été décrites depuis (Guinebretière *et al.*, 2008).

Malgré ces différences, souvent phénotypiques, utilisées pour classer les souches du groupe *Bacillus cereus*, les contours de chaque espèce ne sont pas toujours nets. La différence entre *B. cereus* et *Bt*, par exemple, résiderait dans la présence, chez le second, d'au moins un plasmide portant les gènes nécessaires à la production de toxine Cry. Une souche de *B. thuringiensis* ayant perdu son plasmide Cry aurait alors changé d'espèce ? De même, puisque les déterminants de *B. anthracis* sont plasmidiques, qu'advient-il d'une souche de *B. cereus* ayant acquis les plasmides de virulence *pXO1* et *pXO2* (Bahl and Rosenberg, 2010; Brézillon *et al.*, 2015; Hu *et al.*, 2009) ? Plus généralement, les phylogénies moléculaires du groupe *B. cereus*, qu'elles se basent sur l'AFLP, le MLEE, le MLST, les SNPs ou l'ARNr 16S, peinent à séparer les espèces (exemples en Figs. 1 et 2). Certains auteurs proposent ainsi de définir une espèce regroupant *B. anthracis*, *B. cereus* et *Bt*, espèce qui contiendrait alors des écotypes, pathotypes et thermotypes différents (Guinebretière *et al.*, 2008; Maughan and Van der Auwera, 2011). Ce qui reviendrait à considérer *B. cereus* comme une espèce hautement adaptative, capable, en un minimum de remaniements génomiques, de coloniser des niches aussi variées que le sol, la rhizosphère, les invertébrés et les mammifères.

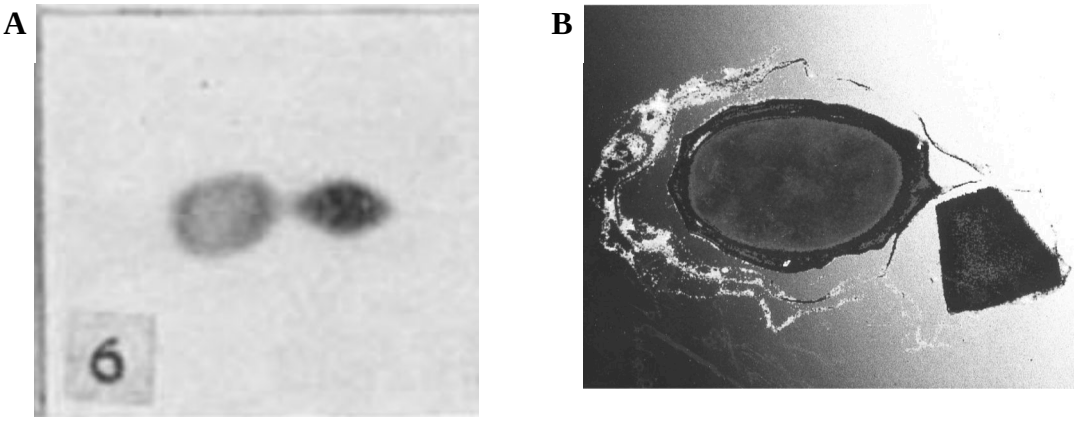


Fig. 3 : Spore de *Bacillus thuringiensis* et son cristal.

A. Première observation rapportée de la spore de *B. thuringiensis* et son cristal associé, en microscopie optique (Hannay, 1953). **B.** Observation en microscopie électronique à transmission (Institut Pasteur).

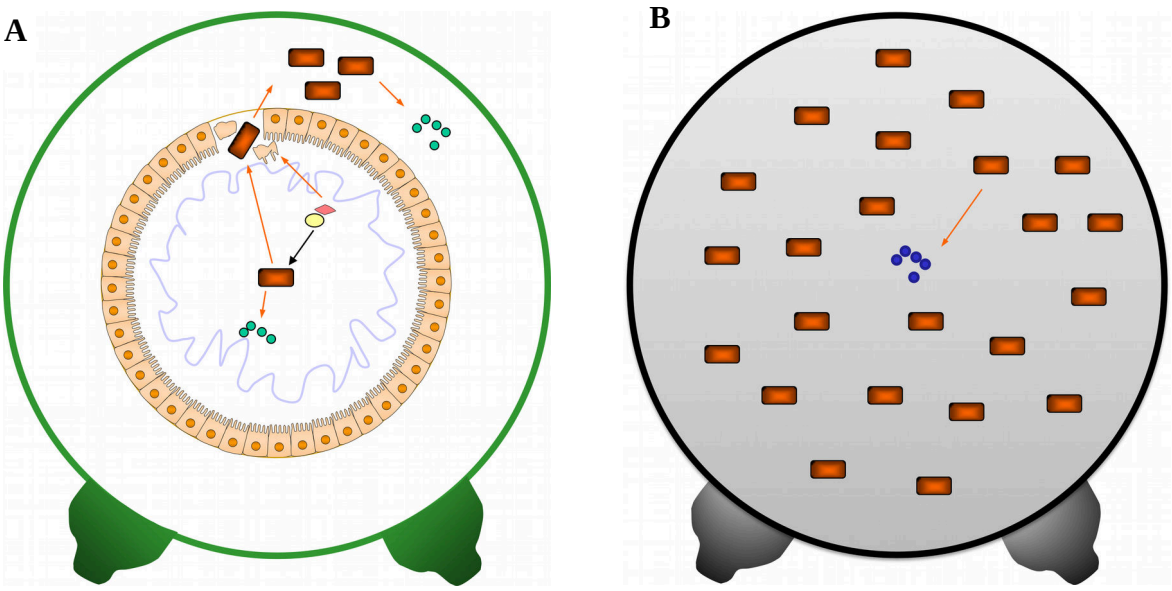


Fig. 4 : Schéma de l'infection d'une larve sensible par *Bacillus thuringiensis*.

A. Après ingestion, les spores et cristaux se retrouvent dans le tube digestif d'une larve vivante. Les toxines Cry déstabilisent la barrière intestinale pendant que les spores germent pour donner des cellules de *B. thuringiensis* végétatives. Ces dernières vont produire une batterie de facteurs de virulence leur permettant d'échapper à l'immunité de l'hôte et de dégrader ses tissus. Lorsque l'épithélium intestinal est détruit, *B. thuringiensis* envahit l'hémocœle. **B.** Une fois la larve morte, *Bt* produit un deuxième jeu d'enzymes lui permettant d'utiliser le cadavre comme source de nutriments, retardant ainsi la sporulation. Figures adaptées de B. Raymond.

Une autre question récurrente est celle de l'origine du groupe *Bacillus cereus*. Une analyse fonctionnelle des génomes du groupe révèle une forte représentation de gènes codant pour des protéases et des protéines du métabolisme et du transport des acides aminés (Ivanova *et al.*, 2003). Cela semble suggérer que l'ancêtre commun aux membres du groupe assimilerait préférentiellement des protéines et acides aminés, régime typique des pathogènes.

2 Le cycle infectieux de *Bacillus thuringiensis*

Bt, comme déjà mentionné, est une bactérie entomopathogène capable d'infecter un large spectre d'insectes, majoritairement de l'ordre des coléoptères, diptères et lépidoptères (van Frankenhuyzen, 2009). Cette pathogénie repose grandement sur la présence dans la souche d'une toxine Cry ou Cyt, produit d'un gène *cry* (plasmidique) ou *cyt* (Deng *et al.*, 2014; Palma *et al.*, 2014; Vachon *et al.*, 2012). Au cours de sa croissance, une population de *Bt* produit des cristaux (Fig. 3). La formation de ces cristaux est souvent concomitante avec la sporulation, mais est parfois plus précoce. Ces produits sont en fait un assemblage cristallin de toxines Cry ou Cyt, et se retrouvent généralement associés à la surface des spores, ce qui leur vaut l'appellation de cristaux parasporaux.

2.1 Destruction de la barrière intestinale et invasion des tissus

Lorsqu'une larve sensible ingère un mélange de spores de *Bt* et de cristaux, le pH élevé du tractus de la larve dissocie le cristal et libère les toxines le composant (Fig. 4) (Clark, 1999; Hannay, 1953; Vachon *et al.*, 2012). Ces toxines libres vont être maturées par clivage protéolytique, dans l'intestin, et ainsi prendre leur conformation active. Les toxines actives vont se lier à des récepteurs membranaires spécifiques de l'épithélium intestinal de l'hôte et y former des pores non-spécifiques. Ces pores, perméables aux ions, acides aminés et sucres, induisent la lyse de certaines cellules épithéliales et la paralysie de l'intestin. Ces événements déstabilisent la barrière intestinale, ralentissant le transit intestinal, ce qui favorise à son tour la germination des spores. Les cellules végétatives peuvent alors commencer à se diviser et à coloniser la surface de l'épithélium intestinal (Ramarao *et al.*, 2012; Salamitou *et al.*, 2000, C. Nielsen-LeRoux, communication personnelle). Les bactéries végétatives sécrètent un cocktail d'enzymes dégradatives (phospholipases, protéases, hémolysines et entérotoxines) sous le

contrôle du régulateur de *quorum sensing* PlcR, ce qui va finir de rompre l'intégrité de l'intestin de l'hôte (Agaisse *et al.*, 1999; Gohar *et al.*, 2002, 2008; Salamitou *et al.*, 2000). Notons que pour certains insectes, l'action de la seule toxine Cry est létale, alors que pour d'autres, la combinaison de la toxine Cry et de *B. thuringiensis* est nécessaire pour induire la mort (Heimpel and Angus, 1959). *Galleria mellonella*, lépidoptère souvent utilisé comme hôte modèle, appartient notamment à cette seconde catégorie (Li *et al.*, 1987; Ramarao *et al.*, 2012).

Dans l'intestin de l'hôte, *B. thuringiensis* fait face à un environnement hostile et ne peut pas se contenter de détruire les tissus. Il doit également se défendre contre les peptides anti-microbiens et le microbiote de la larve. Pour survivre, *B. thuringiensis* met en place plusieurs stratégies, dont la modification de sa paroi (cible des peptides anti-microbiens), la sécrétion de protéases dégradant les molécules anti-bactériennes de l'hôte et la production de ses propres molécules microbicides (Abi Khattar *et al.*, 2009; Dalhammar and Steiner, 1984; de la Fuente-Salcido *et al.*, 2013).

Une fois la barrière intestinale rompue, *B. thuringiensis* envahit l'hémocœle de l'insecte (compartiment renfermant l'hémolymphe, liquide circulatoire équivalant au sang et à la lymphe des vertébrés), où il va rencontrer d'autres acteurs de l'immunité comme les phagocytes, les systèmes de coagulation et la cascade phenoloxidase (Little *et al.*, 2005).

2.2 Nécrotrophisme

Une fois la larve morte, *B. thuringiensis* cesse progressivement de produire ses facteurs de virulence et se met, sous le contrôle d'un second régulateur de *quorum sensing*, NprR, à sécréter un nouveau jeu d'enzymes impliquées dans la dégradation d'autres substrats de l'hôte : des protéases, estérases, lipases, chitinases et une chitosanase, entre autres (Dubois *et al.*, 2012; Verplaetse *et al.*, 2015). Ces enzymes permettent probablement à *B. thuringiensis* d'utiliser le cadavre comme nouvelle source de nutriments, et ainsi de retarder l'entrée en sporulation. Ainsi on peut retrouver des cellules végétatives de *B. thuringiensis* plus d'une semaine après la mort de la larve (Gélis-Jeanvoine *et al.*, observations non publiées).

Lors de cette phase de son cycle infectieux, *B. thuringiensis* produit également un petit métabolite secondaire, la kurstakine, qui sera abordé au chapitre IV de cette introduction.

2.3 Dissémination

Lorsqu'un cadavre d'insecte finit par ne plus rien contenir d'assimilable par *B. thuringiensis*, ce dernier doit, pour survivre, trouver une nouvelle source de nutriments. Or il a été observé que dans un cadavre d'hôte infecté pris à des temps tardifs, la population de *B. thuringiensis* était différenciée en plusieurs sous-populations : une engagée en sporulation (ou déjà sporulée), une restant nécrotrophe, et une troisième non-caractérisée (Dubois *et al.*, 2012; Prasertphon *et al.*, 1973; Verplaetse *et al.*, 2015). La sporulation s'accompagnant, pour *B. thuringiensis*, de la production du cristal parasporal, nous supposons que le complexe spore-cristal retrouvé dans le cadavre constitue le vecteur d'une prochaine infection qui permettra à *B. thuringiensis* de coloniser une nouvelle niche. Les populations restant à l'état végétatif pourraient permettre à *B. thuringiensis*, en cas d'apport de nouveaux nutriments, de se développer plus rapidement qu'une population de spores qui devrait traverser toute la germination auparavant. Ce choix pourrait s'avérer stratégique dans un écosystème complexe en présence de compétiteurs.

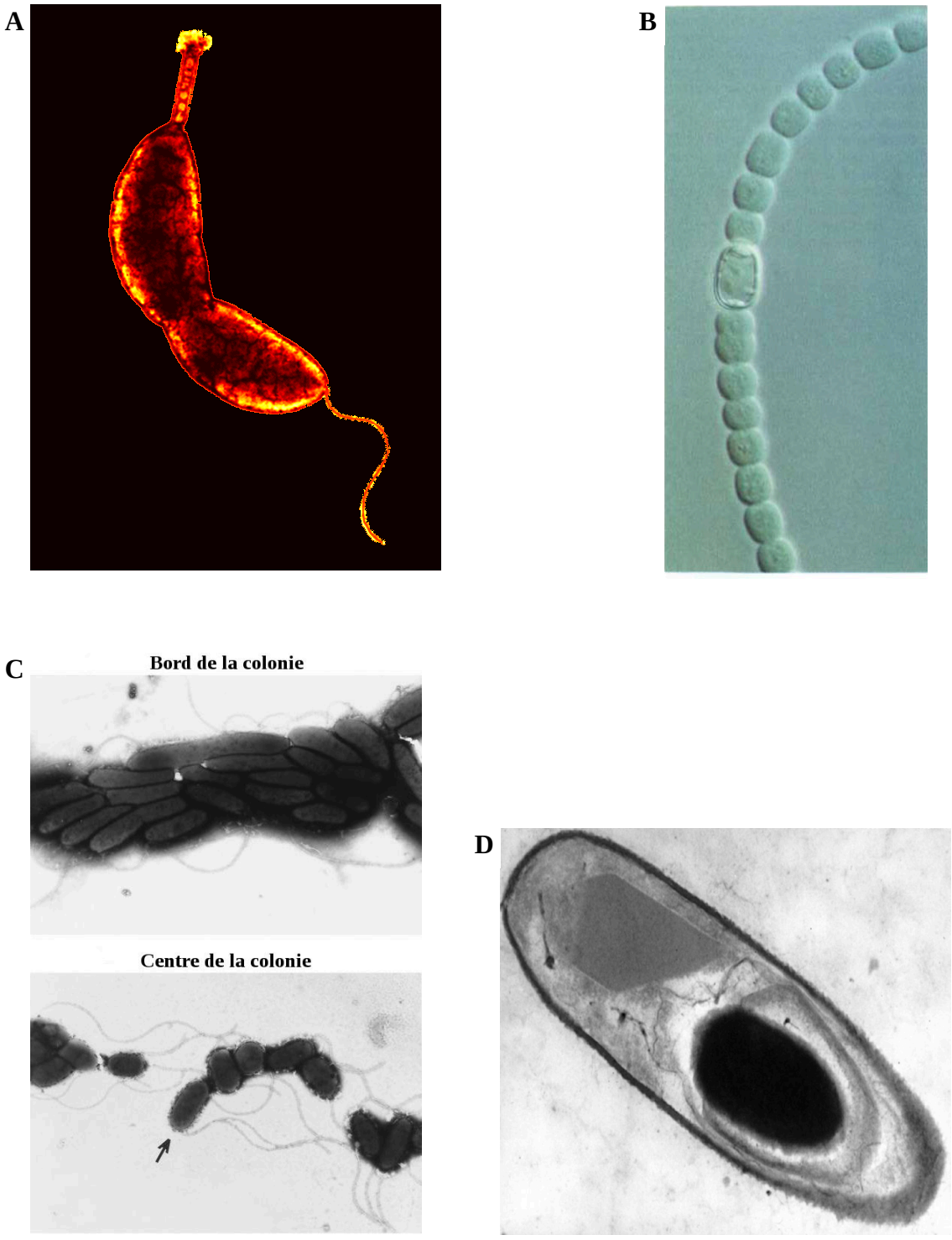


Fig. 5 : Exemples de différenciations bactériennes.

A. *Caulobacter crescentus* en train de se diviser en une cellule flagellée et une cellule à pédicule (Yves Brun). **B.** Chaînette d'*Anabaena* sp. contenant un hétérocyste (James A. Shapiro). **C.** Cellules de *Pseudomonas aeruginosa* swarmeuses (bord de colonie) et non-swarmeuses (centre) (M. Haranur Rashid). **D.** Cellule de *B. thuringiensis* contenant une préspore et un cristal parasporal (Institut Pasteur).

Chapitre II - différenciations bactériennes

1 Les bactéries aussi se différencient

La différenciation, en biologie, est le processus au cours duquel une cellule change de type. En pratique, chez les eucaryotes multi-cellulaires, la différenciation transforme souvent une cellule moins spécialisée en cellule plus spécialisée. Ce phénomène est connu depuis longtemps chez les mammifères et chez d'autres eucaryotes, mais avait peu de sens en parlant de bactéries qu'on pensait strictement uni-cellulaires. Cependant, dès 1988 l'idée est proposée de considérer les populations bactériennes comme des organismes multi-cellulaires, partant précisément de l'observation qu'elles semblent pouvoir se différencier (Shapiro, 1988, 1998). Pour ancienne qu'elle soit, cette idée continue à faire son chemin et la différenciation bactérienne reste un sujet en plein essor. Plusieurs organismes modèles sont sur le devant de la scène, souvent pour des raisons pratiques (de facilité de culture et de manipulation génétique notamment). On notera pêle-mêle *Caulobacter crescentus* et sa division asymétrique donnant naissance à une cellule fille flagellée et une à pédicule, l'hétérocyste des cyanobactéries, le phénomène de mobilité par swarming, ou encore la sporulation et la formation du cristal parasporal de *B. thuringiensis* (Fig. 5). Ces différenciations ont une implication importante : les bactéries doivent être, d'une façon ou d'une autre, capables de faire un choix entre plusieurs modes de développement. Les populations bactériennes ne seraient donc pas de simples amas de clones, mais bien des organismes différenciés. On peut alors qualifier de différenciation bactérienne un processus au cours duquel une partie, plus rarement la totalité, d'une population de bactéries passe d'un mode de développement A à un mode de développement B. La différenciation permet ainsi aux bactéries de s'adapter à leur environnement.

B. thuringiensis est capable de se différencier en fonction de son environnement et au cours de la croissance de sa population. Sa différenciation la plus connue est sûrement la sporulation (Fig. 5D). Mais une population de *B. thuringiensis* est également capable de se différencier pour former un biofilm sur milieu solide (de type colonie) ou à l'interface air – liquide (de type pellicule). Il a également été montré récemment qu'au cours de l'infection d'une larve d'insecte, *B. thuringiensis* se différencie en plusieurs sous-populations : virulente, nécrotrophe, sporulante et indéfinie (Fig. 6) (Verplaetse *et al.*, 2015).

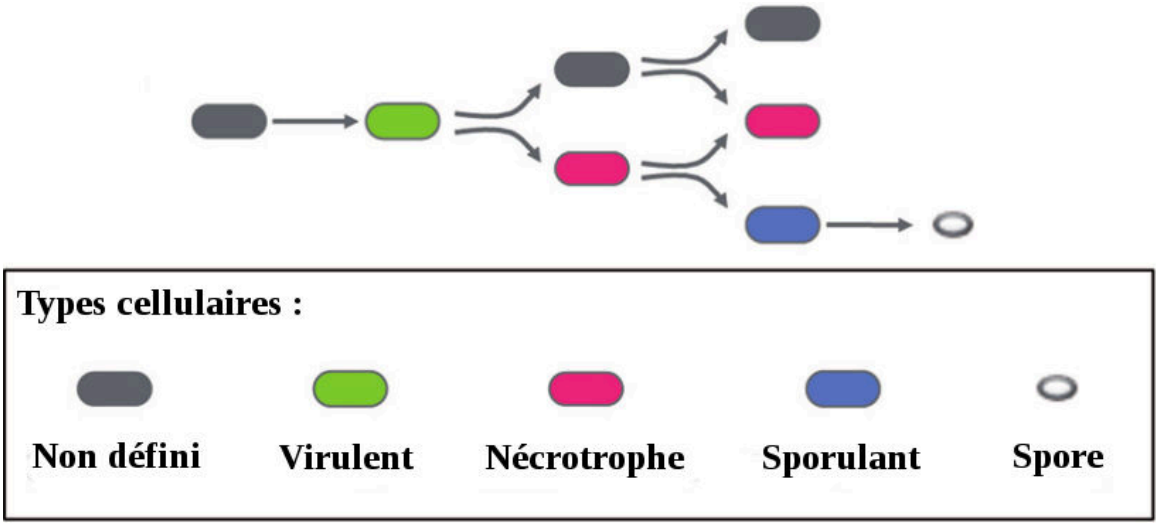


Fig. 6 : différenciations au cours du cycle infectieux de *Bacillus thuringiensis* dans *Galleria mellonella*.

Au cours de son cycle infectieux dans l'hôte, la population de *B. thuringiensis* se différencie en plusieurs sous-populations coexistantes exprimant des régulons différents. Figure adaptée de Verplaetse *et al.* (2015).

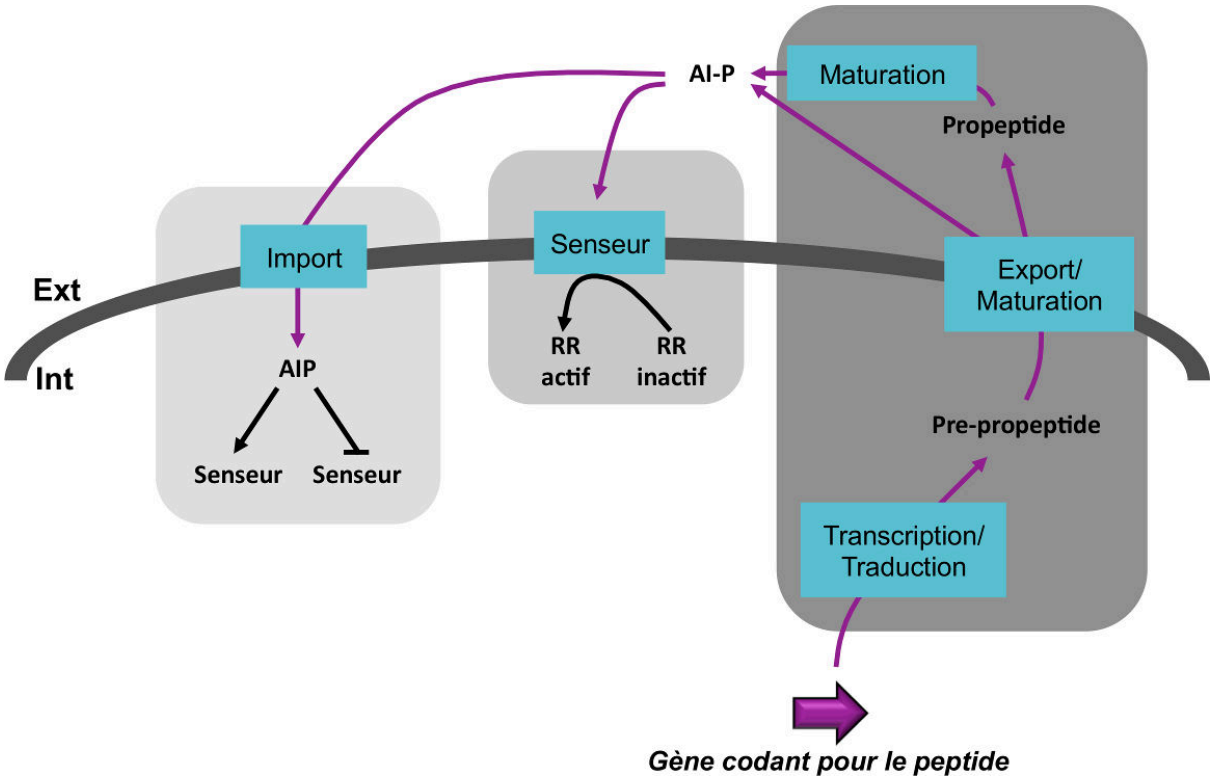


Fig. 7 : Schéma général des systèmes de *quorum sensing* utilisant des peptides.

Figure adaptée de Slamti *et al.* (2014).

2 Les signaux déclenchant la différenciation

2.1 Les signaux « professionnels » : le *quorum sensing*

Une des solutions employées par les bactéries pour déclencher un comportement dans la population est le *quorum sensing*. Le *quorum sensing* est un processus au cours duquel une molécule signal, ou auto-inducteur, est sécrétée et partagée par la population. Lorsque la densité de cette molécule dépasse un seuil, le *quorum*, une réponse est déclenchée. L'exemple historique bien connu est celui de la bactérie endosymbiotique *Vibrio fischeri* et de son hôte, le calamar *Euprimna scolopes* (Verma and Miyashiro, 2013). Au cours de sa croissance dans l'hôte, la population de *Vibrio fischeri* sécrète une acyl-homosérine lactone qui, une fois le *quorum* atteint, déclenche la transcription de la luciférase, une enzyme produisant des photons. Cette coopération permet au calamar d'émettre, *via* ses symbiotes, de la lumière qui lui servirait à masquer son ombre, le rendant ainsi moins facilement repérable par ses prédateurs.

Les signaux de *quorum sensing* utilisés par les bactéries peuvent être regroupés en trois catégories : les auto-inducteurs de type 1 (AI-1) sont des acyl-homosérine lactone et sont utilisés par les bactéries à Gram négatif, les auto-inducteurs peptidiques (AI-P) sont des oligopeptides utilisés par les bactéries à Gram positif, et les auto-inducteurs de type 2 (AI-2) sont des composés cycliques dérivés de la furanone utilisés par les bactéries à Gram positif et négatif. Dans le cas des AI-P, le peptide est synthétisé à partir d'un gène par transcription et traduction, puis exporté et maturé (Fig. 7). Ensuite, soit le peptide mature se lie à un senseur membranaire, soit il est réimporté dans les cellules. Dans le premier cas, la liaison avec le senseur membranaire, une kinase, va entraîner la phosphorylation d'un régulateur de réponse qui, ainsi activé, va réguler la transcription de gènes. Ce type de systèmes, courant chez les bactéries, est appelé système à deux composants. Dans le second cas, le peptide réimporté se lie directement avec le régulateur associé pour l'activer. C'est le cas pour les régulateurs de la famille RNPP (pour Rap, NprR, PlcR et PrgX) (Rocha-Estrada *et al.*, 2010). Les RNPP de *B. thuringiensis*, notamment PlcR et NprR, détaillés plus loin, ont un rôle prépondérant dans le cycle infectieux, comme déjà évoqué plus tôt dans cette introduction.

2.2 Les messagers secondaires

Dans le *quorum sensing*, l'auto-inducteur est sécrété, et c'est la concentration extra-cellulaire qui est détectée indirectement par les régulateurs. Il existe un autre type de molécule signal : les messagers secondaires. Les plus connus, chez les eucaryotes, sont le Ca^{2+} et l'AMPc. Ce sont de petites molécules cytosoliques qui ont souvent d'autres fonctions que celle de signalisation, mais dont la concentration sert d' « indicateur » à certains régulateurs. Le GTP, par exemple, dont la fonction première est de fournir de l'énergie par hydrolyse, est un cofacteur de CodY qui s'en sert de jauge du potentiel énergétique cytoplasmique (Ratnayake-Lecamwasam *et al.*, 2001).

3 Les régulateurs de la différenciation

3.1 Spo0A : le régulateur en chef

L'adaptation est évidemment une somme de processus complexes, multi-factoriels et redondants, mais Spo0A semble néanmoins y occuper une place prépondérante. Ce régulateur a été découvert chez *B. subtilis*, à travers son implication dans les étapes précoces de la sporulation (Piggot and Coote, 1976), mais on sait maintenant qu'il régule en fait, toujours chez *B. subtilis*, de nombreux processus, dont de nombreuses différenciations : la motilité, la formation de biofilm, le cannibalisme, et évidemment la sporulation (Fujita *et al.*, 2005; Grau *et al.*, 2015; López *et al.*, 2009a; Molle *et al.*, 2003a; Piggot and Hilbert, 2004). Spo0A est également impliqué indirectement dans de nombreux autres processus, notamment parce qu'il contrôle l'expression d'autres acteurs majeurs de la phase stationnaire : σ^E, σ^F, σ^G, SinI et AbrB. Chez *B. thuringiensis*, il agit notamment comme répresseur de la transcription de *plcR* (Lereclus *et al.*, 2000).

Le chapitre suivant de cette introduction est dédié à Spo0A, et on se contentera donc ici de quelques généralités. Spo0A est le produit du gène *spo0A*, transcrit dès la fin de la phase exponentielle de croissance. Après transcription et traduction, Spo0A est activé par phosphorylation. Spo0A~P joue alors son rôle de régulateur en activant et réprimant un grand nombre de promoteurs.

3.2 Les facteurs sigma

Les facteurs sigma (σ) sont de petites protéines (environ 200 résidus) impliquées, avec les sous-unités α, β, β' et δ, dans la formation de l'holo-ARN-polymérase (Moran, 1993). Le facteur σ est la sous-unité qui confère une spécificité à l'ARN-polymérase pour certains promoteurs, de sorte qu'on peut parler du régulon d'un facteur σ. Par exemple, le facteur σ^A (équivalent à σ^{70} chez *Escherichia coli*) est le principal facteur σ des cellules végétatives, et contrôle la transcription des gènes de ménage (gènes nécessaires à la vie végétative d'une bactérie) (Haldenwang, 1995). Son régulon est donc très large : 358 unités transcriptionnelles (UT) chez *B. subtilis* (Ishii, 2001; Sierro *et al.*, 2008). D'autres facteurs σ, intervenant dans des processus plus restreints, ont un régulon plus ciblé : c'est le cas de σ^H, le facteur σ principal de la phase de transition, qui ne contrôle l'expression que de 24 UT, toujours chez *B. subtilis*. La famille des facteurs σ^{ECF} regroupe des facteurs σ dont les régulons sont impliqués dans la réponse à des signaux extra-cellulaires (Souza *et al.*, 2014).

Le groupe *Bacillus cereus* contient un grand nombre de facteurs σ, dont une proportion importante de σ^{ECF} (Schmidt *et al.*, 2011). Le génome de *B. thuringiensis* 407 contient quant à lui 29 gènes annotés comme facteurs σ, dont 28 chromosomiques et un plasmidique. Sur ces 29, 10 sont prédits comme σ^{ECF} et 7 comme σ de transition et de sporulation. Si on accepte que la majorité des facteurs σ, et *a fortiori* les σ^{ECF}, est dédiée à l'adaptation, on peut supposer que le grand nombre de facteurs σ de *B. thuringiensis* reflète son adaptation à une niche, l'insecte, plus complexe que celle de *B. subtilis*, dont la souche de laboratoire 168 ne comprend que 18 facteurs σ (dont 7 σ^{ECF}). Il est intéressant de noter, en ce sens, que *Bacillus cytotoxicus* NVH 391-98 est une des souches les plus pauvres en facteurs σ, notamment en σ^{ECF} (Schmidt *et al.*, 2011). Cela pourrait, à l'inverse de *B. thuringiensis* 407, être à rapprocher de sa spécialisation de niche.

3.3 CodY

CodY est un régulateur transcriptionnel pléiotrope présent chez de nombreux firmicutes, et bien étudié chez *B. subtilis* (Sonenshein, 2005). Ses cofacteurs, le GTP et les acides aminés branchés, régulent son activité en reflétant les conditions environnementales : lorsque le milieu est riche en nutriments, les concentrations en GTP et en acides aminés branchés sont hautes et activent CodY ; lorsque le milieu s'appauvrit et que les concentrations de ses cofacteurs chutent, CodY devient inactif.

Chez *B. cereus*, il a été montré que le régulon CodY (direct et indirect) est impliqué dans de nombreuses fonctions, et CodY apparaît comme un « interrupteur » permettant l'allumage et l'extinction de plusieurs modes de développement : production de toxines, motilité, formation de biofilm et adoption d'un métabolisme basé sur les acides aminés (Frenzel *et al.*, 2012; Lindbäck *et al.*, 2012). Le mécanisme par lequel CodY régule la production de toxines a été récemment élucidé : CodY active l'import du peptide activateur de PlcR, qui à son tour active l'expression des gènes de virulence (Slamti *et al.*, 2016).

3.4 AbrB

AbrB est un régulateur transcriptionnel présent chez les firmicutes et impliqué dans la répression des gènes de phase stationnaire (de fonctions très variables) pendant la phase exponentielle (Phillips and Strauch, 2002). Comme pour CodY, le lien entre conditions de croissance et activité d'AbrB est marqué, mais dépend cette fois d'un autre régulateur : Spo0A. Dans un milieu riche en nutriments, *abrB* est exprimé sous le contrôle d'un promoteur σ^A-dépendant et s'auto-réprime, maintenant la concentration d'AbrB dans une fourchette précise lui permettant de réprimer son régulon. Mais lorsque le milieu s'appauvrit *spo0A* est transcrit, Spo0A activé, et une des cibles de Spo0A~P est *abrB*, dont il réprime la transcription (Strauch *et al.*, 1990). La répression du régulon AbrB est alors levée.

Chez *B. thuringiensis*, la cinétique observée de la quantité d'AbrB dans les cellules au cours de la croissance suit le modèle établi chez *B. subtilis*, laissant supposer une analogie fonctionnelle (Lozano Goné *et al.*, 2014). En ce sens, le rôle d'AbrB dans la formation de biofilm a été rapporté comme similaire chez *B. thuringiensis* et *B. subtilis* (Fagerlund *et al.*, 2014). Enfin, il a également été montré que le locus *ces*, responsable de la synthèse de la toxine émétique de *B. cereus*, le céréulide, est réprimé par AbrB (Lücking *et al.*, 2009). Cependant le reste du régulon AbrB de *Bt* et *B. cereus* reste inconnu à ce jour, et la présence de 3 paralogues dans le génome de *B. thuringiensis* 407 suggère une richesse de régulon supérieure à celle observée chez *B. subtilis*.

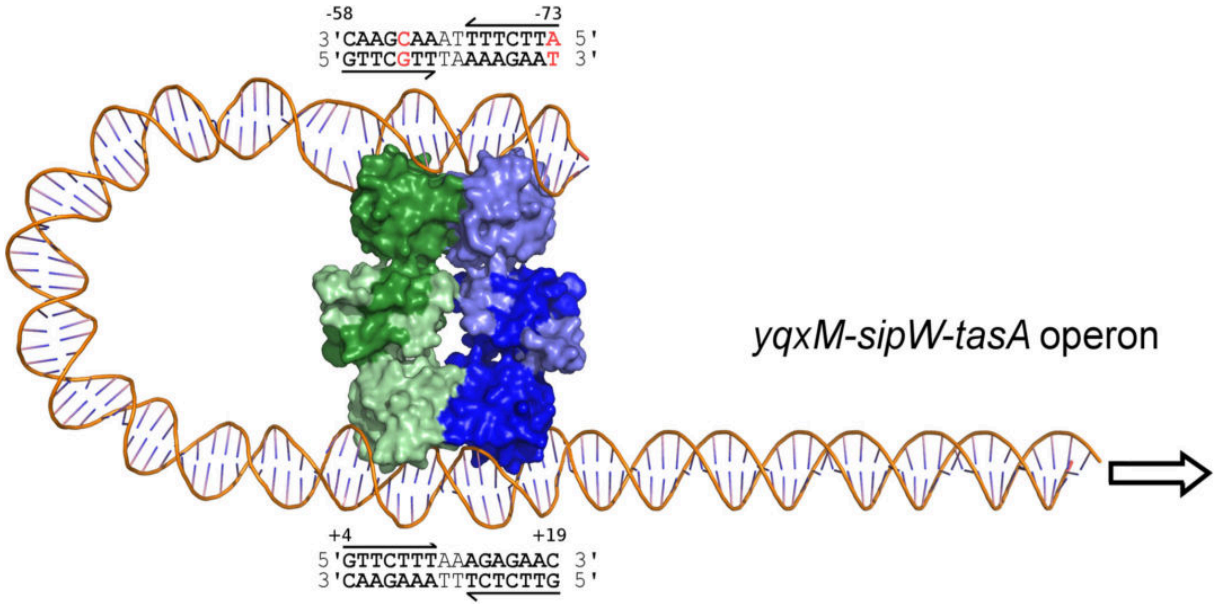


Fig. 8 : Modèle du mécanisme de répression transcriptionnelle de l'opéron *tapA–sipW–tasA* par SinR.

Chaque dimère de SinR est lié à une double-boîte tête-bêche SinR, et la tétramérisation par dimérisation des dimères entraîne le piégeage de la séquence intermédiaire dans une boucle (Newman *et al.*, 2013).

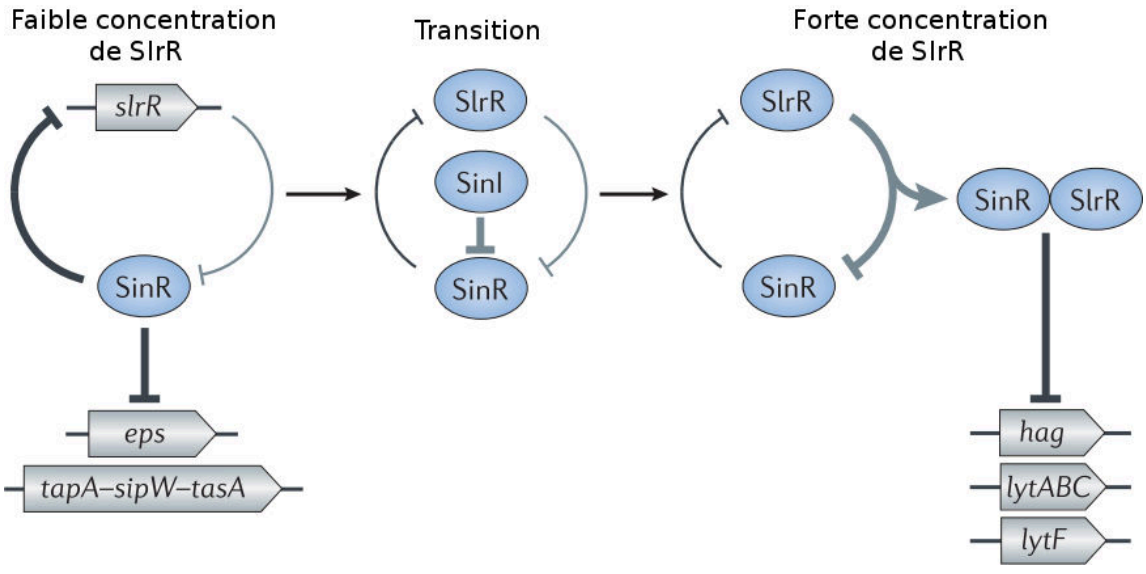


Fig. 9 : Modèle de l'interaction SinR–SlrR.

Les concentrations croissantes de SlrR permettent de réguler finement l'activité de SinR par le biais de régulations génétiques et post-traductionnelles. Figure adaptée de Vlamakis *et al.* (2013).

3.5 SinR

SinR est le régulateur transcriptionnel principal de la formation de biofilms chez les *Bacilli* (Kearns *et al.*, 2005). Son mode d'action a été décrit récemment : un dimère de SinR se fixe sur une première double-boîte SinR, un autre dimère fait de même sur une seconde double-boîte, et la tétramérisation (la dimérisation des dimères) entraîne le piégeage de la séquence d'ADN entre les deux double-boîtes SinR (Fig. 8) (Newman *et al.*, 2013). SinR réprime ainsi notamment la transcription de deux opérons nécessaires à la formation de biofilm (Chu *et al.*, 2006; Kearns *et al.*, 2005).

Le gène *sinR* est en opéron avec le gène codant pour son répresseur SinI. Chez *B. subtilis*, *sinR* est transcrit constitutivement à partir d'un promoteur interne σ^A-dépendant. Dans un milieu riche en nutriments, SinR réprime ainsi l'expression des opérons nécessaires à la formation de biofilm. Lorsque le milieu s'appauvrit, Spo0A~P active la transcription de *sinI* dans une partie de la population, à la fois directement et en levant la répression d'AbrB (Chai *et al.*, 2008; Shafikhani *et al.*, 2002; Strauch, 1995a). Dans ces cellules, SinI antagonise alors SinR en l'immobilisant dans un hétérodimère inactif SinI–SinR (Bai *et al.*, 1993). SinR interagit également avec une autre protéine : SlrR (Fig. 9) (Chu *et al.*, 2008; Kobayashi, 2007). La séquence primaire de SlrR est proche de celle d'une protéine de fusion SinR–SinI. À faible concentration de SlrR, la transcription du gène *slrR* est réprimée par SinR. Puis, lorsque SinR commence à être titré par SinI, la répression est levée, SlrR peut s'accumuler, et un dimère SinR–SlrR peut alors se former. Ce dimère, en plus de titrer SinR comme le fait SinI et donc d'inhiber son activité de répresseur, agit comme répresseur de gènes impliqués dans la mobilité.

Il est intéressant de noter que, bien que SinR soit présent chez *B. subtilis* et *B. thuringiensis*, leurs régulons semblent très différents (Fagerlund *et al.*, 2014). En effet, les deux cibles principales de SinR chez *B. subtilis* sont les opérons *tapA–sipW–tasA* et *epsA–O* (Chu *et al.*, 2006; Kearns *et al.*, 2005). Ces deux loci sont responsables respectivement de la production de la composante protéique fibrillaire de la matrice, et de la synthèse des sucres extra-cellulaires. Des analogues de ces deux loci existent chez *B. thuringiensis*, mais seul l'analogue de *tapA–sipW–tasA* est contrôlé par SinR. De même, les régulons SinR de *B. thuringiensis* et de *B. anthracis* se recouvrent peu (Pflughoeft *et al.*, 2011). Malgré sa conservation chez de nombreux *Bacilli*, SinR apparaît ainsi comme un régulateur de différenciation très spécialisé.

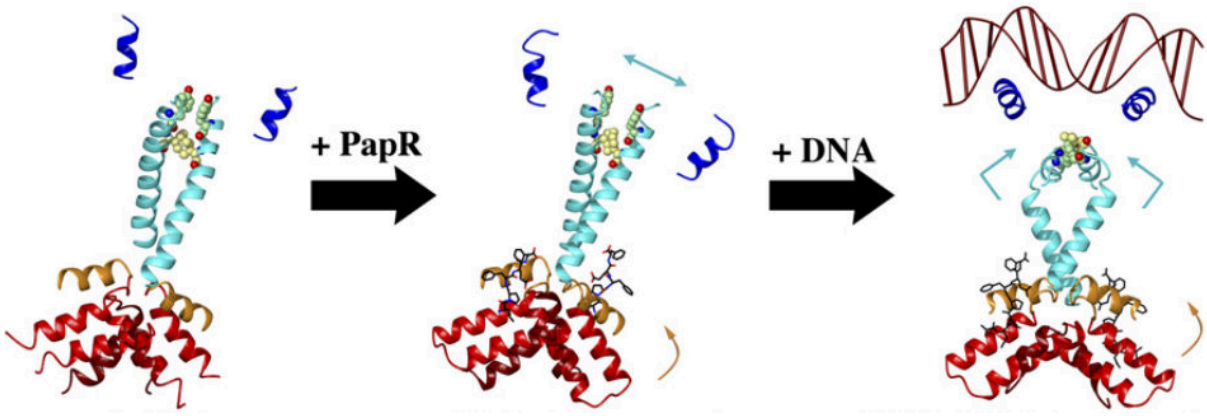


Fig. 10 : Modèle d'activation de PlcR.

La fixation de PapR induit un changement de conformation du dimère de PlcR, notamment des motifs HTH (en bleu foncé) qui prennent l'orientation correcte pour contacter l'ADN (Grenha *et al.*, 2012).

3.6 PlcR

PlcR est un régulateur transcriptionnel présent uniquement dans le groupe *Bacillus cereus*, où il régule l'expression d'un groupe de gènes permettant la virulence dans l'hôte insecte (Agaisse *et al.*, 1999; Gohar *et al.*, 2002, 2008; Salamitou *et al.*, 2000). Le rôle du régulon PlcR a déjà été abordé dans le paragraphe sur le cycle infectieux de *B. thuringiensis*, et on rappellera ici que PlcR active la transcription d'une cinquantaine de gènes codant pour des phospholipases, des protéases, des hémolysines et des entérotoxines. Ce cocktail d'enzymes permet à *B. thuringiensis* d'infecter efficacement son hôte en dégradant ses tissus.

PlcR est un régulateur de *quorum sensing* de la famille des RNPP (Slamti and Lereclus, 2002). La transcription de *plcR* et du gène codant pour son peptide activateur, *papR*, débute en fin de la phase exponentielle (Lereclus *et al.*, 1996). Cette transcription est auto-activée par PlcR et réprimée par Spo0A~P (Lereclus *et al.*, 2000). La liaison de deux PapR à un homodimère de PlcR induit un changement de conformation de ce dernier, de sorte que ses motifs HTH soient dans la bonne orientation pour contacter l'ADN (Fig. 10) (Declerck *et al.*, 2007; Grenha *et al.*, 2012).

Dans le génome de *Bc* se trouvent 3 gènes dont le produit prédit est proche de PlcR : *plcRa*, *plcRb* et *plcRc* (Huillet *et al.*, 2012). Ces paralogues, identiques entre eux à 85 %, présentent environ 29 % d'identité (50 % de similarité) avec PlcR. À ce jour, seul PlcRa a été étudié, et il a été montré que son régulon est impliqué dans la réponse aux stress peroxyde et disulfide (Huillet *et al.*, 2012).

3.7 NprR

NprR est un autre régulateur RNPP important dans le cycle infectieux de *B. thuringiensis*. Comme déjà abordé plus tôt, il est impliqué dans la transition du mode de vie pathogène au mode de vie nécrotrophe (Dubois *et al.*, 2012). Comme PlcR, son régulon contient de nombreux gènes codant pour des enzymes dégradatives : protéases, estérases, lipases, chitinases et une chitosanase. Le régulon NprR est nécessaire pour la survie de *B. thuringiensis* dans le cadavre de son hôte. L'hypothèse proposée par les auteurs est que les enzymes sous le contrôle de NprR servent à exploiter le cadavre comme source de nutriments, permettant ainsi à *B. thuringiensis* de retarder l'entrée en sporulation.

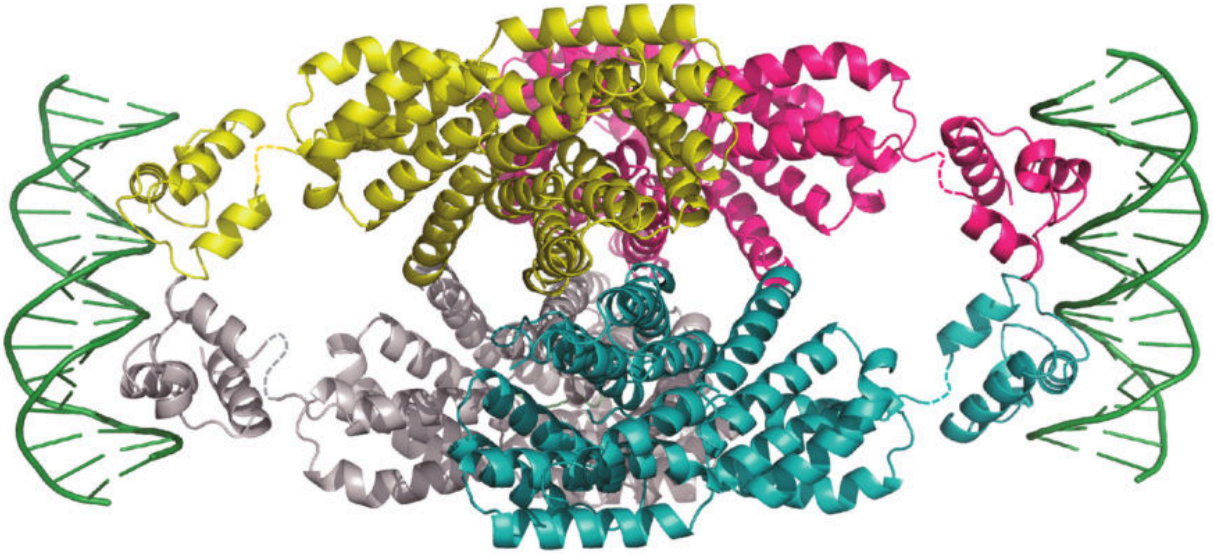


Fig. 11 : Modèle du tétramère de NprR lié à l'ADN.

La structure du cœur de NprR a été résolue, mais les motifs HTH (au contact de l'ADN, deux à droite et deux à gauche) et l'interface avec l'ADN ont été modélisés (Zouhir *et al.*, 2013).

Le gène *nprR* est en opéron avec le gène codant pour son peptide activateur, *nprX* (Dubois *et al.*, 2013). La transcription de l'opéron *nprR–nprX* est réprimée par CodY en phase exponentielle et activée par PlcR~PapR en fin de phase exponentielle. Le contrôle transcriptionnel exercé par PlcR sur *nprR–nprX* permet d'expliquer la transition qui s'opère du régulon PlcR vers le régulon NprR. Entre *nprR* et *nprX* se trouvent également deux promoteurs relais, σ^E- et σ^H-dépendants qui permettent le maintien d'une expression prolongée d'*nprX*. Sous forme *apo* (non-liée à NprX), NprR est homodimérique, et la liaison d'NprX entraîne sa tétramérisation, forme sous laquelle les motifs HTH seraient correctement orientés pour contacter l'ADN (Fig. 11) (Zouhir *et al.*, 2013). Il a été démontré récemment au laboratoire que l'homodimère de NprR jouait également un rôle de régulateur, post-traductionnel cette fois, en empêchant partiellement la phosphorylation de Spo0A et donc la sporulation (Perchat *et al.*, article soumis).

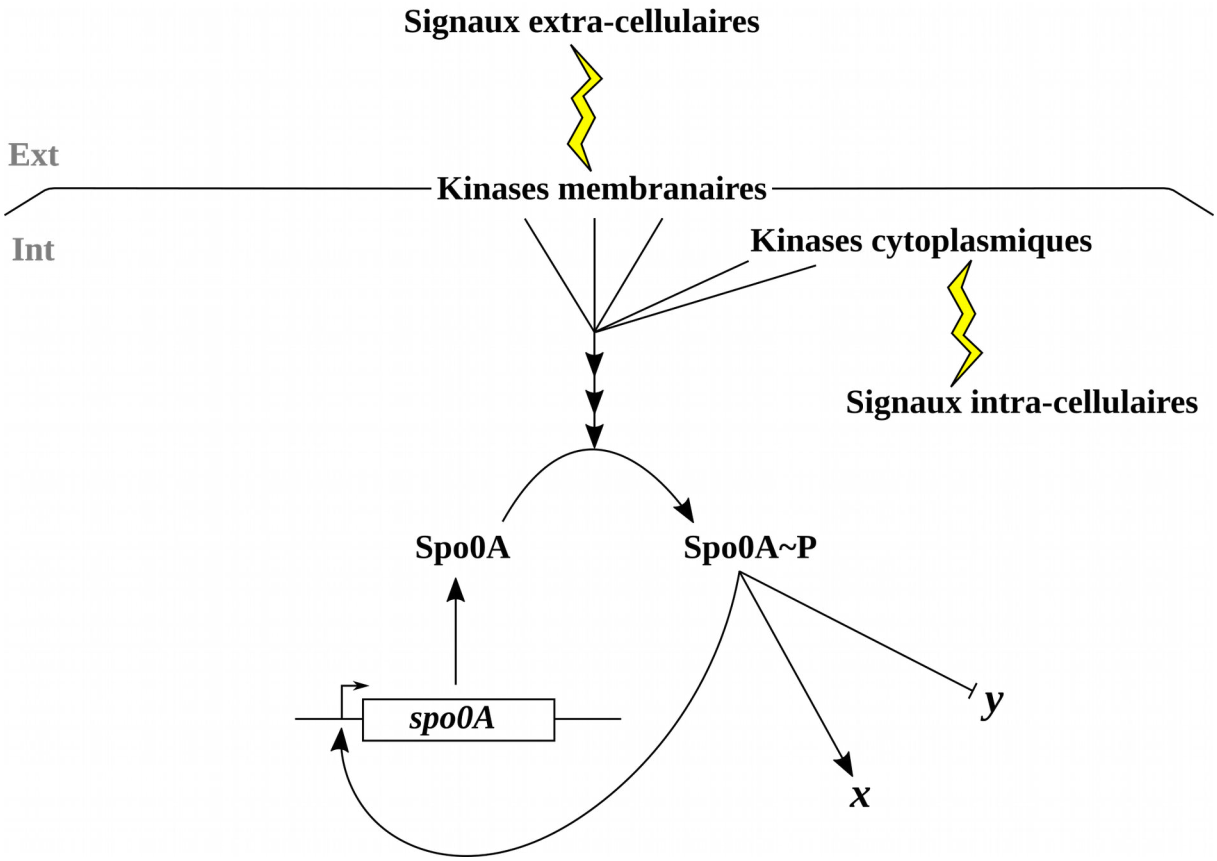


Fig. 12 : Schéma général de l'activation de Spo0A.

Les symboles jaunes représentent les signaux activateurs des kinases. *x* et *y* symbolisent l'ensemble des gènes activés et réprimés, respectivement, par Spo0A~P. Ext : extérieur. Int : intérieur.

Chapitre III - Spo0A, le régulateur clé de la différenciation

Spo0A est un régulateur transcriptionnel pléiotrope présent chez toutes les bactéries sporulantes connues (Galperin, 2013). Il y est qualifié de régulateur global (« master regulator ») car ses régulons et le nombre de différenciations sous son contrôle sont conséquents. Spo0A a été découvert par plusieurs équipes au début des années 1970 comme étant responsable d'une étape précoce de la sporulation chez *B. subtilis* (Piggot and Coote, 1976). Depuis, une somme importante de travail a été accomplie pour aboutir au modèle (complexe) des voies Spo0A-dépendantes tel qu'on le connaît actuellement, et de nombreux travaux viennent encore régulièrement affiner ce modèle. Dans ce modèle, Spo0A fait, comme il avait été initialement pressenti, office d'initiateur, d'interrupteur principal. Nous tenterons, dans ce chapitre, d'aborder les régulations transcriptionnelles de *spo0A* et post-traductionnelles de Spo0A, aussi bien que le rôle de Spo0A activé. La majorité des travaux évoqués ici concerne *B. subtilis*, avec quelques allusions à *Clostridium difficile*, et un dernier paragraphe est consacré aux données disponibles chez *B. thuringiensis*.

1 Schéma général

Avant d'aborder le détail de la régulation de et par Spo0A, il est nécessaire d'avoir une vue d'ensemble du système, cela à cause des nombreux rétro-contrôles qu'il contient (Fig. 12). En premier lieu, le gène *spo0A* est transcrit et traduit en Spo0A. Spo0A, sous cette forme, est inactif. Lorsque des signaux activent les kinases dites de sporulation, celles-ci déclenchent une cascade de phosphorylation qui aboutit à la phosphorylation de Spo0A. Spo0A~P active et réprime la transcription de nombreux gènes possédant une « boîte Spo0A » (site de fixation de Spo0A sur l'ADN). Sur ce schéma est notamment figurée l'auto-activation de *spo0A* par Spo0A~P.

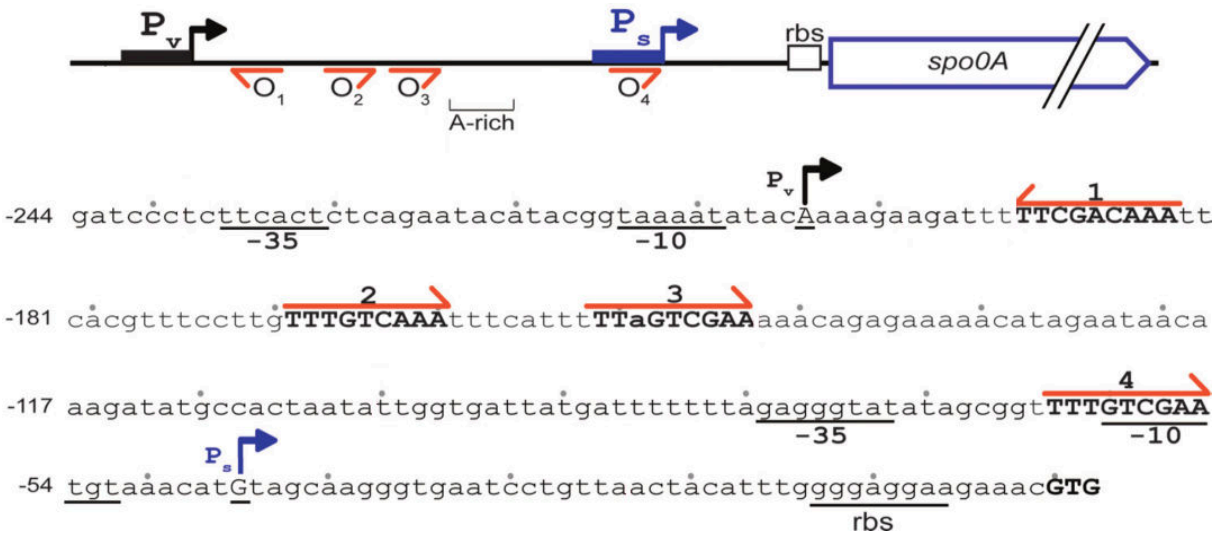


Fig. 13 : *spo0A* et son promoteur chez *Bacillus subtilis* 168.

En haut, une représentation schématique de *spo0A* et son promoteur. Les deux promoteurs y sont représentés : Pv, promoteur végétatif et Ps, promoteur de sporulation. Les quatre boîtes Spo0A sont figurées en orange et nommées O_1–O_4. En dessous, la séquence annotée du promoteur. Figure adaptée de Chastanet & Losick (2011).

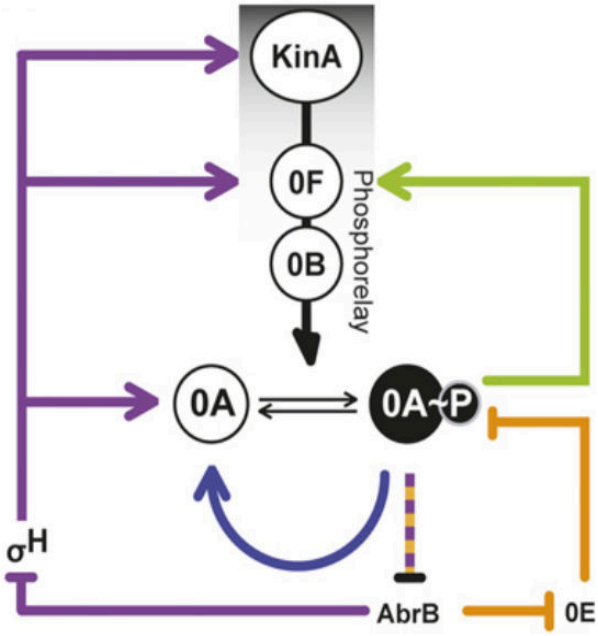


Fig. 14 : Les boucles de rétro-contrôle impliquées dans la transcription et l'activation de Spo0A.

Spo0A~P exerce quatre boucles de rétro-contrôle sur lui-même : 1) En activant sa propre transcription, 2) en activant la transcription de *spo0F*, un élément du phosphorelais, 3) en réprimant la transcription d'*abrB*, levant ainsi la répression de σ^H qui active la transcription de *spo0A*, et 4) toujours en réprimant la transcription d'*abrB*, levant ainsi la répression de *spo0E* dont le produit déphosphoryle Spo0A~P (Chastanet *et al.*, 2010).

2 Régulation transcriptionnelle de *spo0A*

Chez *B. subtilis*, la régulation de la transcription de *spo0A* est complexe et fait intervenir deux facteurs σ, trois boîtes Spo0A et quatre boucles d'auto-régulation (Figs. 13 et 14) (Chastanet and Losick, 2011; Chastanet *et al.*, 2010; Chibazakura *et al.*, 1991).

Au cours de la phase exponentielle de croissance, *spo0A* est transcrit à partir d'un promoteur σ^A-dépendant (Pv, pour promoteur végétatif) (Chibazakura *et al.*, 1991). Cette transcription est faible, et le transcrit est fortement structuré en 5', gênant la traduction, mais cela permet tout de même le maintien d'un pool d'environ 2000 molécules de Spo0A par cellule (Chastanet and Losick, 2011; Chastanet *et al.*, 2010).

À la fin de la phase exponentielle, l'activation de Spo0A permet la levée de la répression de σ^H par AbrB, car Spo0A~P réprime la transcription d'*abrB* (Strauch, 1995b). S'opérerait alors un changement de promoteur pour *spo0A* : Pv se retrouverait inhibé par Spo0A~P qui s'est fixé à la boîte O_1 alors que la transcription à partir de Ps (pour promoteur de sporulation), σ^H-dépendant, pourrait démarrer. On note que σ^H active également la transcription d'éléments nécessaires à la phosphorylation de Spo0A. La transcription à partir de Ps serait aussi activée par Spo0A~P fixé sur la boîte O_3, et réprimée par la boîte O_2. Le transcrit démarrant de Ps ne possède pas la région très structurée en 5' de celui démarrant de Pv, aussi les auteurs supposent que sa traduction se fait plus facilement. L'activation de Spo0A, en plus de promouvoir la transcription de *spo0A* à partir de Ps et de permettre l'expression de σ^H, active également, toujours à travers la levée de répression par AbrB, l'expression d'une phosphatase qui déphosphoryle (et donc désactive) Spo0A. Enfin, Spo0A~P augmente sa propre activation en promouvant la transcription d'un gène codant pour un élément nécessaire à sa phosphorylation, Spo0F. Ces quatre boucles d'auto-régulation sont résumées en figure 14. Cette régulation complexe qui se met en place dès la fin de phase exponentielle permet l'augmentation progressive du nombre de molécules de Spo0A par cellule, jusqu'à dépasser les 20 000.

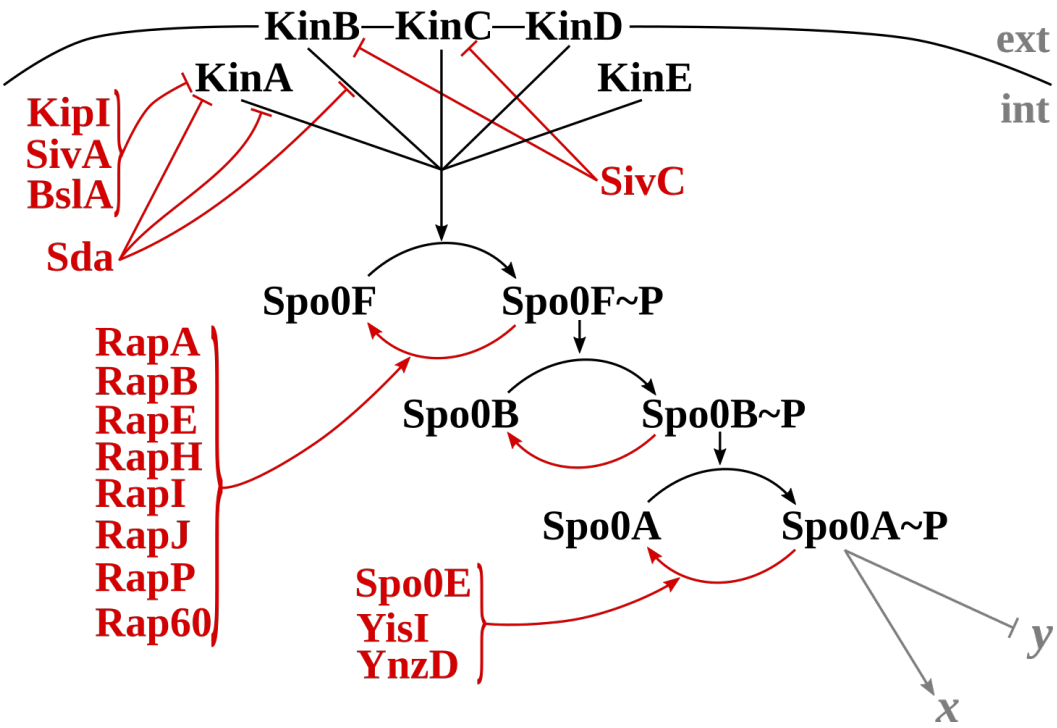


Fig. 15 : Schéma du modèle de phosphorelais chez *Bacillus subtilis*.

En noir, les éléments du phosphorelais à proprement parler. En rouge, les différents inhibiteurs du phosphorelais. En gris, *x* et *y* symbolisent l'ensemble des gènes activés et réprimés, respectivement, par Spo0A~P. Ext : extérieur. Int : intérieur.

3 Régulation post-traductionnelle de Spo0A

Après transcription et traduction, Spo0A se trouve sous sa forme inactive, c'est-à-dire trop peu affine pour l'ADN pour promouvoir la transcription (Trach *et al.*, 1991). La phosphorylation sur un résidu aspartate de sa partie N-terminale induit un changement de conformation de Spo0A, le rendant beaucoup plus affin pour l'ADN, et donc actif (Lewis *et al.*, 2000; Trach *et al.*, 1991; Zhao *et al.*, 2002). Mais contrairement aux systèmes à deux composants connus, Spo0A n'est pas un régulateur de réponse classique auquel est associé une kinase. Le système est plus complexe et fait intervenir *a minima*, chez *B. subtilis*, cinq kinases, cinq anti-kinases, huit phosphatases et deux protéines intermédiaires, un système baptisé « phosphorelais » (Fig. 15) (Burbulys *et al.*, 1991).

3.1 Les kinases

Les éléments les plus amont du phosphorelais sont les kinases KinA-E (Jiang *et al.*, 2000). Ces histidine-kinases partagent une surface d'interaction avec Spo0F, leur substrat principal. Il a été montré que KinC et KinD sont également capables de phosphoryler directement Spo0A, bien que Spo0F semble lui être préféré (Jiang *et al.*, 2000). KinD aurait également un double-rôle de kinase et phosphatase, ce qui lui conférerait la capacité de retarder l'entrée en sporulation (Aguilar *et al.*, 2010).

Un enjeu majeur est la compréhension des signaux modulant l'activité de ces kinases, longtemps restés mystérieux. Depuis quelques années cependant, un certain nombre de signaux et de partenaires ont été montrés comme capables d'induire l'auto-activation ou l'inhibition de certaines de ces kinases, mais la liste reste probablement très incomplète. L'auto-phosphorylation de KinA et le phospho-transfert de KinA à Spo0F est notamment inhibé par Sda, une petite protéine dont la production est induite par les dommages à l'ADN et le blocage de la réplication (Burkholder *et al.*, 2001; Rowland *et al.*, 2004). L'auto-phosphorylation de KinA est également inhibée par KipI, une petite protéine dont l'expression est régulée par la concentration en azote, et par SivA et BslA, deux petites protéines dont l'expression est activée par Spo0A~P (Garti-Levi *et al.*, 2013; Wang *et al.*, 1997). L'activité de KinB requiert la présence de deux partenaires membranaires : une lipoprotéine, KapB, et une protéine à six hélices trans-membranaires, KbaA (Dartois *et al.*, 1996, 1997; Trach and Hoch, 1993). KinC serait quant à elle activée en réponse à un efflux de potassium généré,

entre autres, par la surfactine, un lipopeptide détaillé dans le chapitre IV de cette introduction (López *et al.*, 2009b). L'auto-activation de KinB et KinC est inhibée par SivC, une petite protéine sous le contrôle transcriptionnel de Spo0A~P (Garti-Levi *et al.*, 2013). KinD serait activée par des signaux aussi divers que le mélange glycérol-manganèse, les sucres de la matrice extra-cellulaire et la pression osmotique (Aguilar *et al.*, 2010; Rubinstein *et al.*, 2012; Shemesh and Chai, 2013). Rien ne semble connu, pour l'instant, sur la modulation de l'activité de KinE. On note que parmi ces quelques signaux connus, peu sont liés directement à l'environnement. Comme l'entrée en sporulation repose malgré tout sur les conditions environnementales, il faut probablement imaginer que certains signaux captés par les kinases sont indirects, et non pas une (ou des) molécule(s) clé du milieu extérieur. C'est, par exemple, le cas de CodY qui se sert de la concentration intra-cellulaire de GTP et d'acides aminés branchés comme indicateurs du pool énergétique de la cellule, reflet de la disponibilité en nutriments dans le milieu de croissance. Beaucoup de travaux sont encore nécessaires pour relier ce que l'on connaît des conditions de sporulation à l'activité des kinases.

En plus de cette batterie de signaux, on note que les gènes codant pour les kinases sont soumis à une régulation transcriptionnelle fine, complexifiant encore le modèle (Fujita and Sadaie, 1998; Molle *et al.*, 2003b; Strauch, 1995b; Tojo *et al.*, 2013).

3.2 Le relais et les phosphatases

Chez *B. subtilis*, les kinases auto-phosphorylées transmettent leur phosphate à Spo0F, régulateur de réponse, qui le transfère à Spo0B, phosphotransférase, qui le transfère finalement à Spo0A (Trach *et al.*, 1991). Il a été proposé que l'existence de ce phosphorelais, plutôt qu'un classique système à deux composants, se justifie par l'apport de bruit dans le signal (Chastanet *et al.*, 2010). Ce bruit permettrait, en réponse à des signaux induisant la sporulation, de ne pas déclencher une réponse homogène dans toute la population, un choix qui pourrait ne pas être le plus judicieux.

Une possibilité concrète pour bruiter le signal est l'intervention de phosphatases dans le phosphorelais (Fig. 15). En effet, il a été montré que de nombreuses protéines sont capables de déphosphoryler Spo0F et Spo0A. Les plus connues sont sûrement les Rap, dont huit ont été décrites comme ayant Spo0F pour substrat : RapA, RapB, RapE, RapH, RapI, RapJ, RapP et Rap60 (Boguslawski *et al.*, 2015; Parashar *et al.*, 2011, 2013a, 2013b; Pottathil and Lazazzera, 2003; Smits *et al.*, 2007). Les systèmes Rap60 et RapP ont été découverts

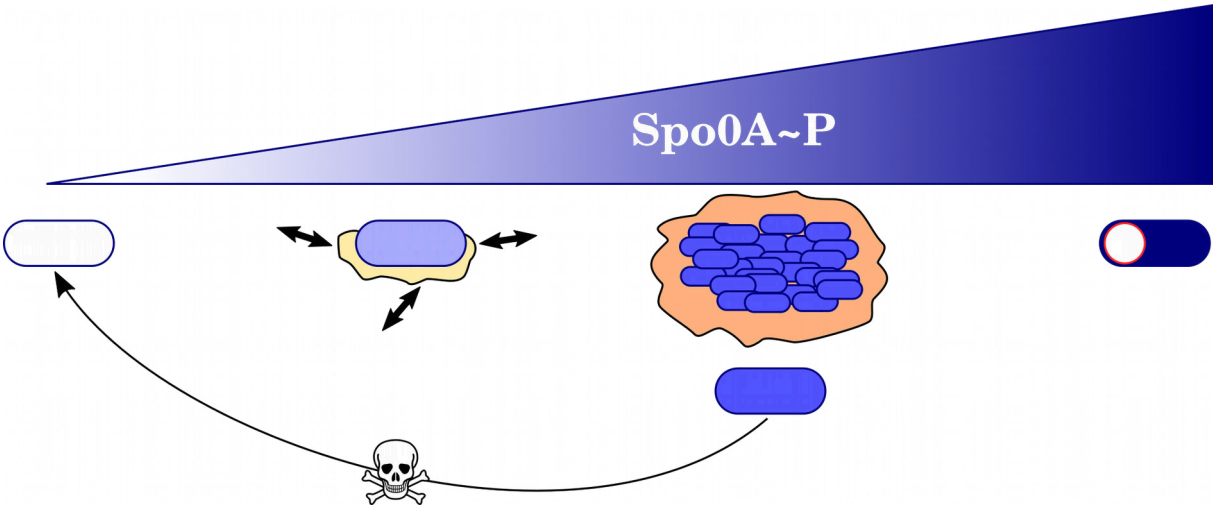


Fig. 16 : Accumulation de Spo0A~P et différenciations de *Bacillus subtilis*.

L'accumulation de Spo0A~P déclenche successivement plusieurs différenciations : à faible concentration, le sliding, à concentration intermédiaire, la formation de biofilm et le cannibalisme, et à forte concentration, la sporulation.

récemment sur des plasmides de *B. subtilis* (Boguslawski *et al.*, 2015; Parashar *et al.*, 2013b). Ces Rap seraient capables de se lier à Spo0F pour empêcher sa phosphorylation par Spo0A, et à Spo0F~P pour le déphosphoryler et/ou bloquer le transfert de phosphate vers Spo0B (Parashar *et al.*, 2011). Les Rap, comme évoqué plus tôt dans cette introduction, appartiennent à la famille RNPP et leur activité est donc régulée par des peptides de *quorum sensing*, les Phr (Pottathil and Lazazzera, 2003). La liaison entre une Rap et son Phr entraîne un changement de confirmation rendant la Rap incapable d'interagir avec son partenaire, dans le cas des RapA, B, E, H, I, J, P et 60, Spo0F, et est donc un événement favorable à la sporulation (Gallego del Sol and Marina, 2013). D'autres phosphatases ont pour substrat Spo0A~P directement : c'est le cas de Spo0E, YnzD et YisI, trois protéines analogues (Perego, 2001).

Chez *C. difficile*, l'absence de Spo0F et Spo0B a rapidement mené à suggérer que les histidine-kinases orphelines (sans régulateur de réponse associé) présentes dans son génome seraient en fait les kinases de sporulation, et auraient pour substrat direct Spo0A. Cela a été prouvé plus tard pour trois des cinq kinases orphelines (Underwood *et al.*, 2009). Les signaux modulant potentiellement l'activité de ces kinases n'ont, à ce jour, pas encore été décrits. Cependant le génome de *C. difficile* contient un orthologue de *kipI*, qui pourrait donc jouer le même rôle d'anti-kinase que chez *B. subtilis* (Edwards and McBride, 2014).

4 Les voies de différenciation régulées par Spo0A~P

4.1 Une question de concentration

Comme introduit plus tôt, Spo0A~P contrôle plusieurs voies de différenciation, notamment la formation de biofilm et la sporulation. La question se pose alors : comment un même régulateur peut être responsable de processus non-concomitants, faisant intervenir des régulons différents ? Chez *B. subtilis*, il a été d'abord pressenti puis montré que la clé réside dans l'accumulation progressive de Spo0A~P dans la cellule (Fujita *et al.*, 2005; Kovács, 2015). C'est en fait l'affinité variable de Spo0A~P pour les différentes boîtes Spo0A du génome qui, couplée à des systèmes de phosphorylation-déphosphorylation redondants exprimés de façon hétérogène et à des boucles de rétro-contrôle complexes, vus plus tôt, qui permet d'échelonner le régulon Spo0A dans le temps. Cette architecture complexe fait de Spo0A un régulateur central dans la différenciation qui contrôle, chez *B. subtilis*, la mobilité par sliding, le cannibalisme, la formation de biofilm et la sporulation. Le modèle qui prévaut actuellement veut que l'accumulation de Spo0A~P progressive active un premier régulon

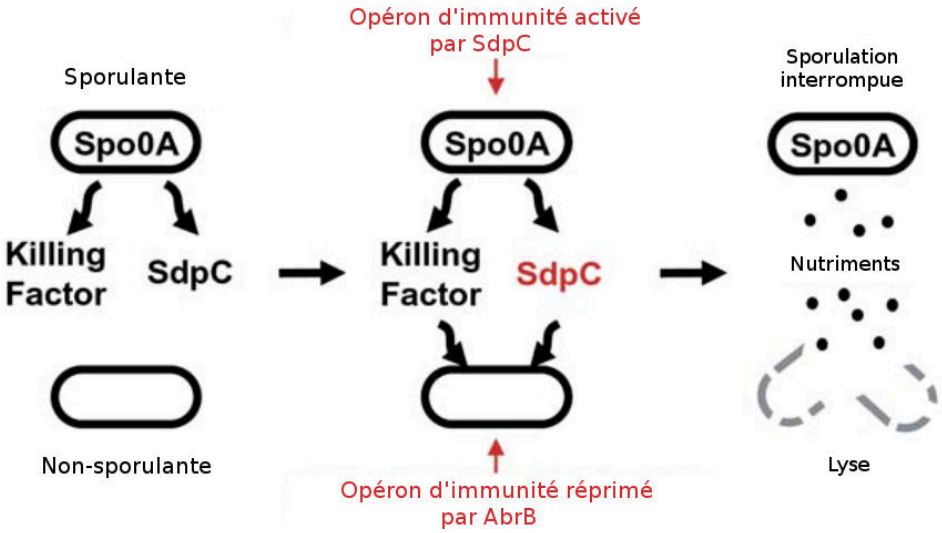


Fig. 17 : Modèle du cannibalisme chez *Bacillus subtilis*.

Les cellules sporulantes sécrètent le « killing factor » et SdpC, deux facteurs induisant la lyse des cellules. Ces mêmes cellules s'immunisent contre l'action de ces facteurs, ce qui n'est pas le cas des cellules non-sporulantes. Ces dernières lysent donc sous l'action des deux facteurs sécrétés. L'hypothèse proposée par les auteurs pour expliquer ce phénomène est que la lyse des cellules non-sporulantes permet de fournir des nutriments aux cellules sporulantes et ainsi de retarder la sporulation. Figure adaptée d'Ellermeier *et al.* (2006).

permettant notamment la mobilité par sliding, puis le cannibalisme et la formation de biofilm, et enfin la sporulation (Fig. 16) (Fujita *et al.*, 2005; Kovács, 2015; López and Kolter, 2010). Chez *C. difficile* et d'autres pathogènes sporulants, Spo0A régule également la production de toxines.

4.2 La mobilité par « sliding »

La mobilité par sliding, contrairement à toutes les autres (swimming, swarming, gliding et twitching), ne requiert aucun appendice extra-cellulaire et est dite « passive » (Jarrell and McBride, 2008). Ce processus ferait en revanche intervenir la surfactine, les sucres extra-cellulaires de la matrice, BslA et la synthèse de lipides *de novo* (Grau *et al.*, 2015). Une seule étude lie à ce jour Spo0A~P à la mobilité par sliding, et démontre que ce mode de développement est activé par de faibles concentrations de Spo0A phosphorylé par KinB et KinC (Grau *et al.*, 2015).

4.3 Le cannibalisme

Le cannibalisme a été décrit chez *B. subtilis* comme étant une différenciation précédant la sporulation. En tout début de phase de sporulation, la population se diviserait en fait en deux : une sous-population qui entre en sporulation, induisant la lyse de la seconde sous-population (Ellermeier *et al.*, 2006; González-Pastor *et al.*, 2003).

Le cannibalisme fait intervenir un peptide ressemblant à un antibiotique, Skf (pour « sporulation killing factor »), la toxine protéique SdpC (pour « sporulation delaying protein »), et une protéine membranaire d'immunisation, SdpI (Fig. 17). Les cellules se dirigeant vers la sporulation contiennent à ce moment une concentration intermédiaire de Spo0A~P, suffisante pour activer la transcription des opérons *sdp* et *skf* (Fujita *et al.*, 2005). Initialement, la transcription d'SdpI est réprimée par AbrB. Dans les cellules se dirigeant vers la sporulation, la sécrétion d'SdpC déclenche une boucle d'auto-activation qui, en plus de la levée de répression d'AbrB par Spo0A~P, permet l'accumulation de SdpI dans la membrane et donc l'immunisation. Dans le reste de la sporulation, la présence de Spo0A~P à de faibles concentrations, non-seulement ne déclenche pas la boucle d'accumulation d'SdpI, mais permet en plus à AbrB de réprimer fortement sa transcription, ce qui la laisse sensible à la lyse par SdpC et Skf. La lyse de cette sous-population permettrait de fournir des nutriments aux

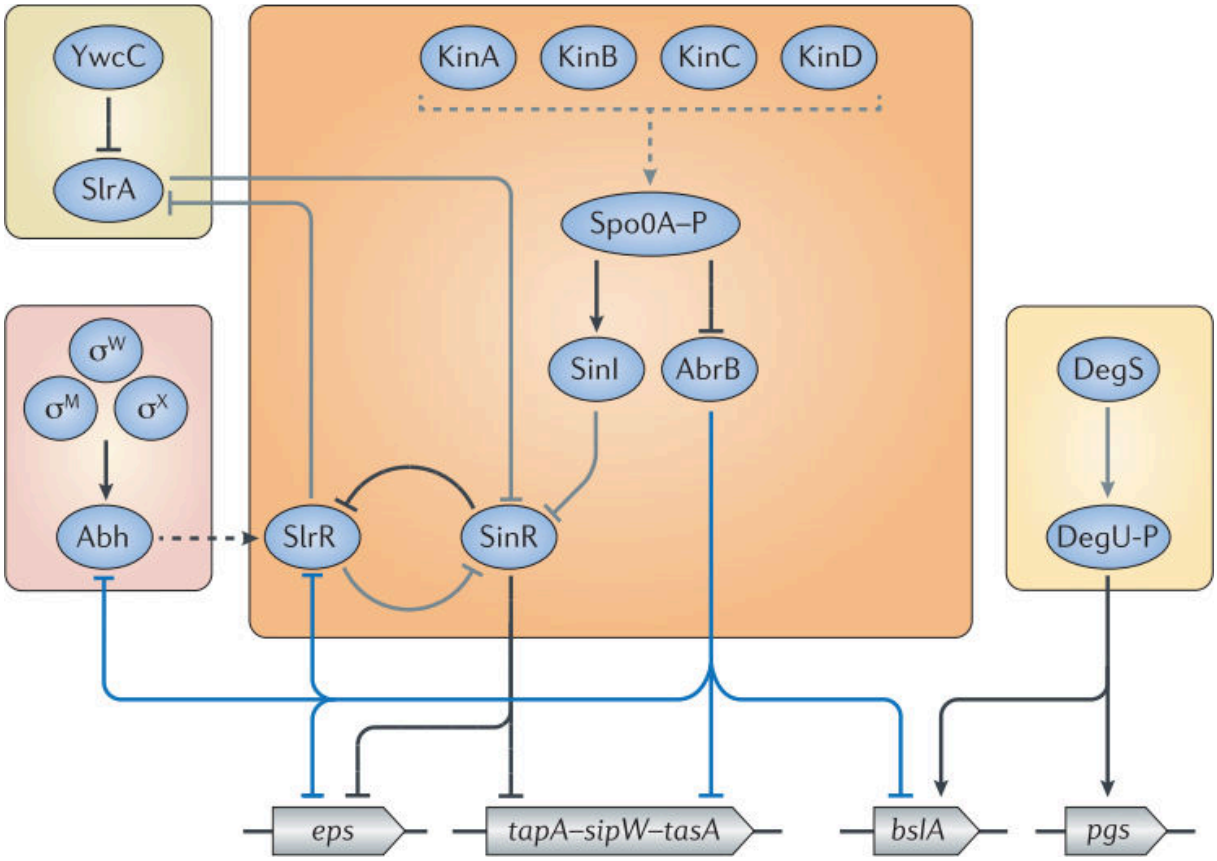


Fig. 18 : Modèle de régulation de la formation de biofilm chez *Bacillus subtilis*.

Spo0A~P, en régulant la transcription de *sinI* et *abrB*, contrôle l'expression des gènes impliqués dans la formation de biofilm. Figure adaptée de Vlamakis *et al.* (2013).

cellules restantes, ce qui permettrait de retarder l'accumulation de Spo0A~P et donc, *in fine*, la sporulation.

Ce système est donc régulé à la fois directement par Spo0A~P qui, à concentration intermédiaire, active la transcription des opérons *sdp* et *skf*, et indirectement *via* AbrB qui réprime la transcription d'*sdpI*.

4.4 La formation de biofilm

Le biofilm peut être décrit, de façon très générique, comme une communauté microbienne partageant une matrice extra-cellulaire et se développant sur une surface. Il est également souvent cité comme un mode de résistance aux stress environnementaux, car il permet à la communauté de supporter des conditions qui induiraient la lyse de cellules isolées. Chez *B. subtilis*, le biofilm est une structure complexe formée par différenciation sous le contrôle, non-exclusif mais principal, de Spo0A~P (Fig. 18) (Vlamakis *et al.*, 2013). Deux types majeurs de biofilms sont étudiés chez cet organisme : celui se développant à l'interface air–agar (la colonie) et la pellicule se développant à l'interface air–liquide. Dans ces deux modèles, trois loci ont été identifiés comme essentiels à la structuration du biofilm : les opérons *epsA–O* et *tapA–sipW–tasA* et le gène *bslA*. Les gènes de l'opéron *epsA–O* sont impliqués dans la synthèse et la sécrétion des exopolysaccharides de la matrice. L'opéron *tapA–sipW–tasA* permet la synthèse (*tasA*), la maturation (*sipW*) et la translocation (*tapA*) de la composante fibrillaire protéique de la matrice extra-cellulaire (Branda *et al.*, 2006; Romero *et al.*, 2010, 2011; Stöver and Driks, 1999a, 1999b; Tjalsma *et al.*, 1998). BslA confère une hydrophobie de surface aux cellules du biofilm.

La formation de biofilm, comme le cannibalisme, est déclenchée par des concentrations intermédiaires de Spo0A~P, et les deux différenciations se déclenchent d'ailleurs de façon concomitante dans la même sous-population (López *et al.*, 2009c). Spo0A~P à concentration intermédiaire permet la formation de biofilm par deux mécanismes : l'activation de la transcription de *sinI* et la répression de la transcription d'*abrB*. Comme introduit dans le paragraphe sur SinR, la forte transcription de *sinI* sous le contrôle de Spo0A~P dans une sous-population de cellules permet, *via* l'interaction SinI–SinR, de lever la répression que SinR exerce sur les opérons *tapA–sipW–tasA* et *epsA–O* (Bai *et al.*, 1993; Chai *et al.*, 2008; Chu *et al.*, 2006; Kearns *et al.*, 2005; Shafikhani *et al.*, 2002; Strauch, 1995a). Spo0A~P lève également la répression d'AbrB sur son régulon, qui comprend notamment les

opérons *tapA–sipW–tasA* et *epsA–O* et le gène *bslA* (Chu *et al.*, 2008; Hamon *et al.*, 2004; Verhamme *et al.*, 2009). Ces levées de répressions orchestrées par Spo0A~P, à travers SinR et AbrB, permettent finalement le démarrage de la formation de biofilm.

C. difficile est également capable de former des biofilms dans différents systèmes, mais ces biofilms sont dits immergés, car ils se forment au fond du contenant. De plus, le génome de *C. difficile* ne contient ni opéron *tapA–sipW–tasA* ni opéron *eps*, et il faut donc supposer que le mécanisme de formation du biofilm diffère fortement de celui décrit chez *B. subtilis*. Peu de données sont disponibles à ce jour sur la régulation de la formation du biofilm de *C. difficile*. Il a cependant été montré que Spo0A, le di-GMPc et un système de quorum sensing de type AI-2 étaient impliqués dans la transition de la vie planctonique vers le biofilm (Đapa *et al.*, 2013; Dawson *et al.*, 2012; Soutourina *et al.*, 2013).

4.5 La sporulation

La sporulation bactérienne, déjà abordée dans cette introduction, est un processus au cours duquel une cellule mère végétative forme puis libère une endospore. L'endospore est une forme de résistance métaboliquement inactive qui peut, lorsque les conditions sont réunies, germer pour donner une nouvelle cellule végétative. Seules quatres classes du phylum des *Firmicutes* contiennent des espèces sporulantes : les *Bacilli*, les *Clostridia*, les *Erysipelotrichia*, et les *Negativicutes* (Galperin, 2013). La sporulation est un processus très complexe faisant intervenir, directement et indirectement, un très grand nombre de gènes. Cette cascade d'événements a été, et continue d'être étudiée en détail, notamment chez *B. subtilis* (Higgins and Dworkin, 2012; Piggot and Hilbert, 2004). Nous ne détaillerons pas ici la totalité de la voie de sporulation, mais simplement son initiation par Spo0A~P.

L'entrée en sporulation, chez *B. subtilis*, nécessite de fortes concentrations de Spo0A~P (Fujita *et al.*, 2005). Pour déclencher la sporulation, Spo0A~P active l'expression de trois unités transcriptionnelles *spoII* : *spoIIAA–spoIIAB–sigF*, *spoIIGA–sigE–sigG*, et *spoIIE* (Baldus *et al.*, 1994; York *et al.*, 1992). La première permet la synthèse de σ^F, le facteur σ précoce de la préspore. La seconde permet la synthèse et la maturation de σ^E, le facteur σ précoce de la cellule mère. Il permet également, plus tard renforcé par une boucle de rétro-contrôle positif, de synthétiser σ^G, le facteur σ tardif de la préspore. Enfin, SpoIIE participe à l'activation de σ^F. Ces facteurs σ sont évidemment tous nécessaires au processus de sporulation, et de σ^E et σ^F dépend toute la suite de la cascade de signalisation menant à la

formation d'une endospore viable. On note que l'activation des régulons σ^E et σ^F, chez *B. subtilis*, engage la cellule de façon irréversible dans la sporulation (Parker *et al.*, 1996). La nécessité d'accumuler de fortes concentrations de Spo0A~P avant d'entrer en sporulation est souvent interprétée comme une hiérarchisation des réponses au stress, la sporulation étant le dernier recours. En effet, on peut imaginer que le coût évolutif de la sporulation fait « préférer » à *B. subtilis* le changement de niche *via* la mobilité, la formation de biofilm ou même le cannibalisme avant d'en venir à sporuler (Weller and Wu, 2015).

Chez *C. difficile*, espèce également sporulante, la cascade d'événements semble légèrement différer de celle de *B. subtilis*. Comme déjà évoqué plus tôt, le phosphorelais tel que décrit chez *B. subtilis* est absent du génome de *C. difficile*, et Spo0A serait phosphorylé directement par d'autres kinases (Underwood *et al.*, 2009). L'absence d'AbrB laisse également imaginer des divergences majeures avec le modèle *B. subtilis*. Les étapes précoces impliquant σ^E et σ^F seraient relativement conservées entre les deux modèles, alors que les étapes suivantes, ainsi que la communication entre préspore et cellule mère, le seraient nettement moins (Pettit *et al.*, 2014; Saujet *et al.*, 2014). Le modèle global de la voie de sporulation chez *C. difficile* proposé à ce jour repose encore grandement sur celui établi chez *B. subtilis*, mais les données spécifiques à *C. difficile* commencent à abonder et dépeignent un modèle de régulation alternatif (Pettit *et al.*, 2014; Saujet *et al.*, 2014).

4.6 La production de toxines

C. difficile produit deux exotoxines entérotoxiques pour les mammifères : TcdA et TcdB (Awad *et al.*, 2014). Les gènes *tcdA* et *tcdB* sont transcrits à partir d'un promoteur TcdR-dépendant (Mani and Dupuy, 2001). TcdR est un facteur σ^{ECF} codé par le gène éponyme localisé dans le même îlot de pathogénicité que *tcdA* et *tcdB*.

Le rôle de Spo0A~P dans la transcription de tcdA et tcdB est sujet à controverse, mais il semble établi que Spo0A~P module l'expression de ces gènes indirectement, différemment en fonction des souches étudiées (Mackin *et al.*, 2013; Pettit *et al.*, 2014; Rosenbusch *et al.*, 2012).

5 Et chez *Bacillus thuringiensis* ?

Il est très important de noter que la quasi-totalité du modèle de la voie Spo0A a été établi chez *B. subtilis*, et qu'il n'est probablement qu'en partie transposable aux autres espèces, y compris du genre *Bacillus*. Dans les paragraphes précédents, les allusions à *C. difficile* ont permis de souligner les différences entre ces deux modèles. Dans ce paragraphe, nous nous concentrons sur le peu de données disponibles chez *B. thuringiensis*.

Le génome de *B. thuringiensis* 407 contient des gènes prédits pour coder les homologues potentiels de Spo0F, Spo0B, Spo0A, ainsi que les facteurs σ de transition et de sporulation σ^H, σ^E, σ^F, σ^G et σ^K. Trois gènes *sigF* sont même annotés : *sigF2* présente la même organisation opéronique que le *sigF* de *B. subtilis*, alors que *sigF1* et *sigF3* semblent appartenir à d'autres opérons flanqués de gènes phagiques. D'après les données de RNA-seq présentées dans les résultats de cette thèse, les trois gènes sont exprimés, bien qu'à des niveaux différents. Deux gènes *sigE* sont également annotés, un sur le chromosome et l'autre sur le plasmide de 500 kb pBTB_502p. La copie chromosomique présente le même contexte génétique que l'homologue de *B. subtilis*, mais pas la copie plasmidique qui semble être monocistronique et très peu transcrite. Les régulateurs globaux Spo0A~P-dépendants SinR et AbrB sont également présents dans le génome de *B. thuringiensis* 407. L'environnement génétique de l'opéron *sinI–sinR* est similaire, mais pas identique à celui de *B. subtilis*. Pour rappel, les cibles principales de SinR chez *B. subtilis* sont les opérons *tapA–sipW–tasA* et *epsA–O*. Chez *B. thuringiensis*, on retrouve un analogue partiel de *tapA–sipW–tasA* ainsi que deux loci *eps*. Le locus *eps1* ressemble à l'*epsA–O* de *B. subtilis* (Bennaceur, 2013), mais pas *eps2* (Majed, données non-publiées). Quatre copies du gène *abrB* sont annotées, *abrB1* présentant un contexte similaire à l'*abrB* de *B. subtilis*, et *abrB2*, *abrB3* et *abrB4* un contexte totalement différent, à la fois entre eux et de celui de *B. subtilis*. La transcription d'*abrB2* est régulée positivement par PlcRa (Huillet *et al.*, 2012). Ces duplications de facteurs σ et de partenaires essentiels de Spo0A laissent entrevoir des régulations différentes et plus complexes que le modèle établi chez *B. subtilis*. En revanche, les gènes impliqués dans le cannibalisme sont absents du génome de *B. thuringiensis*.

Le promoteur de *spo0A* revêt une importance cruciale, car de sa régulation transcriptionnelle dépend toute la suite de la cascade d'événements. Comme exposé dans un paragraphe précédent, la transcription de *spo0A* a été étudiée en détail chez *B. subtilis*, et s'est révélée complexe. Cependant, à notre connaissance, aucune étude ne semble avoir abordé la

transcription de *spo0A* chez *B. thuringiensis* à ce jour. Pourtant, si la séquence codante du génome de *B. thuringiensis* 407 prédit un produit très proche du Spo0A de *B. subtilis*, les promoteurs ne présentent aucune analogie de séquence. Notamment, une seule boîte Spo0A est retrouvée dans le promoteur *spo0A* de *B. thuringiensis* 407, contre quatre chez *B. subtilis*. Cela laisse, encore une fois, envisager une régulation transcriptionnelle fort différente.

Les seules études abordant Spo0A chez *B. thuringiensis* décrivent chacune une petite partie de son régulon. Il a ainsi été montré que Spo0A réprime la transcription de certains gènes *cry* (Agaisse and Lereclus, 1994; Poncet *et al.*, 1997) et de *plcR* (Lereclus *et al.*, 2000), et active celle de l'opéron *dlt* impliqué dans la D-alanylation des acides téichoïques (Kamar, 2014), de gènes impliqués dans l'accumulation de granules de PHB (Chen *et al.*, 2010), et indirectement, *via* AbrB, des céréulide-synthétases (Lücking *et al.*, 2009). Phénotypiquement, Spo0A a également été montré comme étant impliqué, comme chez *B. subtilis*, dans la formation de biofilm et la sporulation (Fagerlund *et al.*, 2014; Lereclus *et al.*, 1995).

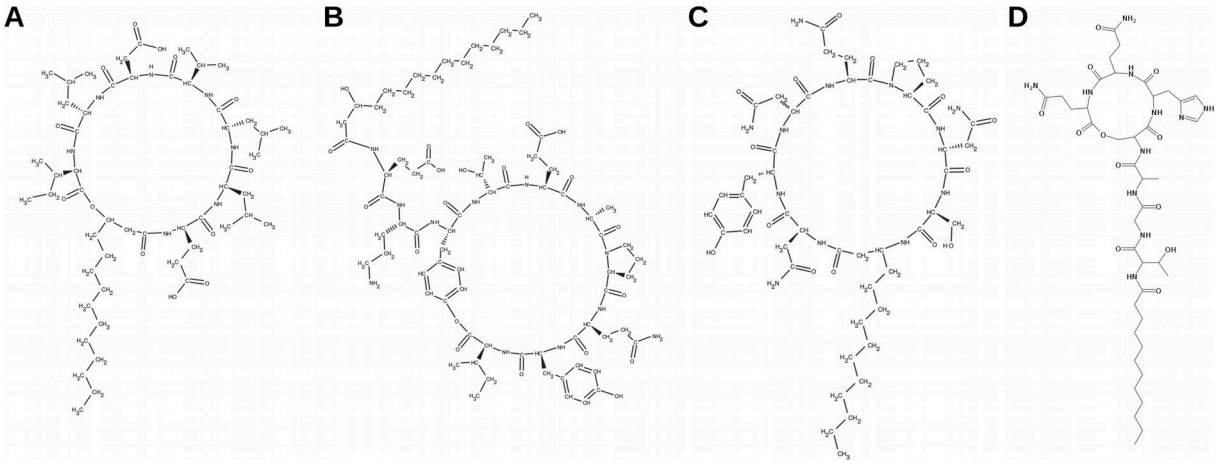


Fig. 19 : Représentants des 4 familles de lipopeptides de *Bacillus*.

La surfactine (A), la fengycine (B), l'iturine (C) et la kurstakine (D).

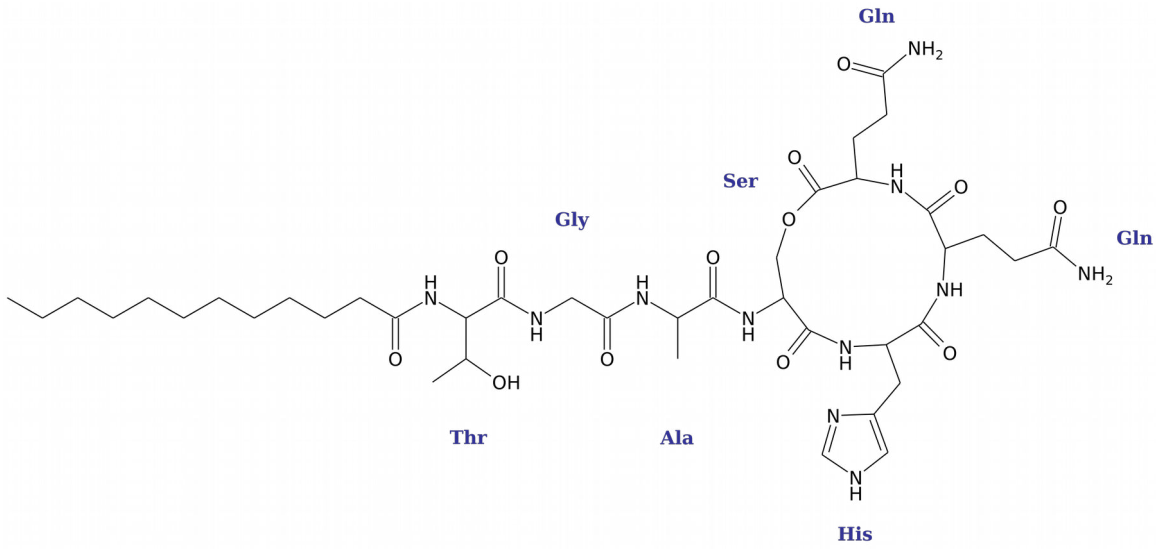


Fig. 20 : Structure annotée de kurstakine cyclique avec acide gras en C_{12} β-hydroxylé.

Chapitre IV - La kurstakine

1 Les lipopeptides de *Bacillus*

Les lipopeptides sont de petits métabolites secondaires produits par certaines espèces des genres *Pseudomonas* et *Bacillus* (Jacques, 2011). Ils sont, comme leur nom l'indique, composés d'une partie peptidique et d'une partie lipidique, ce qui les rend par nature amphiphiles. Cette propriété leur vaut l'appellation de biosurfactants, car ces molécules sont tensioactives et réduisent donc la tension superficielle des solutions aqueuses. Cela les rend également capables d'interagir avec, et de déstabiliser, les membranes biologiques (Inès and Dhouha, 2015).

De nombreux lipopeptides de *Bacillus* ont été décrits, et sont regroupés en quatre familles, selon leur formule : les surfactines, les fengycines, les iturines et la kurstakine (Fig. 19). La famille des surfactines regroupe la surfactine, la pumilacidine, la lichenysine, l'espérine et la bamylocine. La famille des fengycines regroupe la fengycine et la plipastatine. La famille des iturines regroupe l'iturine, la bacillomycine et la mycosubtiline. Enfin, la kurstakine forme une famille distincte, à un seul membre pour l'instant. Ces molécules sont souvent cycliques, leur dernier acide aminé étant lié au premier acide aminé (cycle complet) ou à un acide aminé interne (cycle partiel).

Les lipopeptides de *Bacillus* revêtent un intérêt pour de nombreuses industries : cosmétique, environnementale, phytosanitaire, alimentaire et médicale, entre autres. Ils y sont utilisés comme tensioactifs, émulsifiants, antifongiques, antibactériens, antiviraux, anti-tumoraux, stimulants de la résistance des plantes, *etc* (Cameotra and Makkar, 2004; Cawoy *et al.*, 2013; Kanlayavattanakul and Lourith, 2010; Khong *et al.*, 2012; Nitschke and Costa, 2007). La littérature abonde donc de données biotechnologiques à leur sujet, mais contient nettement moins d'études sur leur rôle *in vivo*. Il a malgré tout été montré que les lipopeptides de *Bacillus* peuvent être impliqués dans la compétition avec d'autres micro-organismes (bactéries, champignons et virus), la motilité, la formation de biofilms et l'induction de la résistance systémique des plantes (Raaijmakers *et al.*, 2010). Leur rôle dans la formation de biofilm est complexe et probablement multi-factoriel. Il pourrait dépendre directement des caractéristiques physico-chimiques de la molécule : selon la partie exposée de la molécule, le lipopeptide pourrait favoriser l'adhésion à des surfaces hydrophiles ou hydrophobes. Les

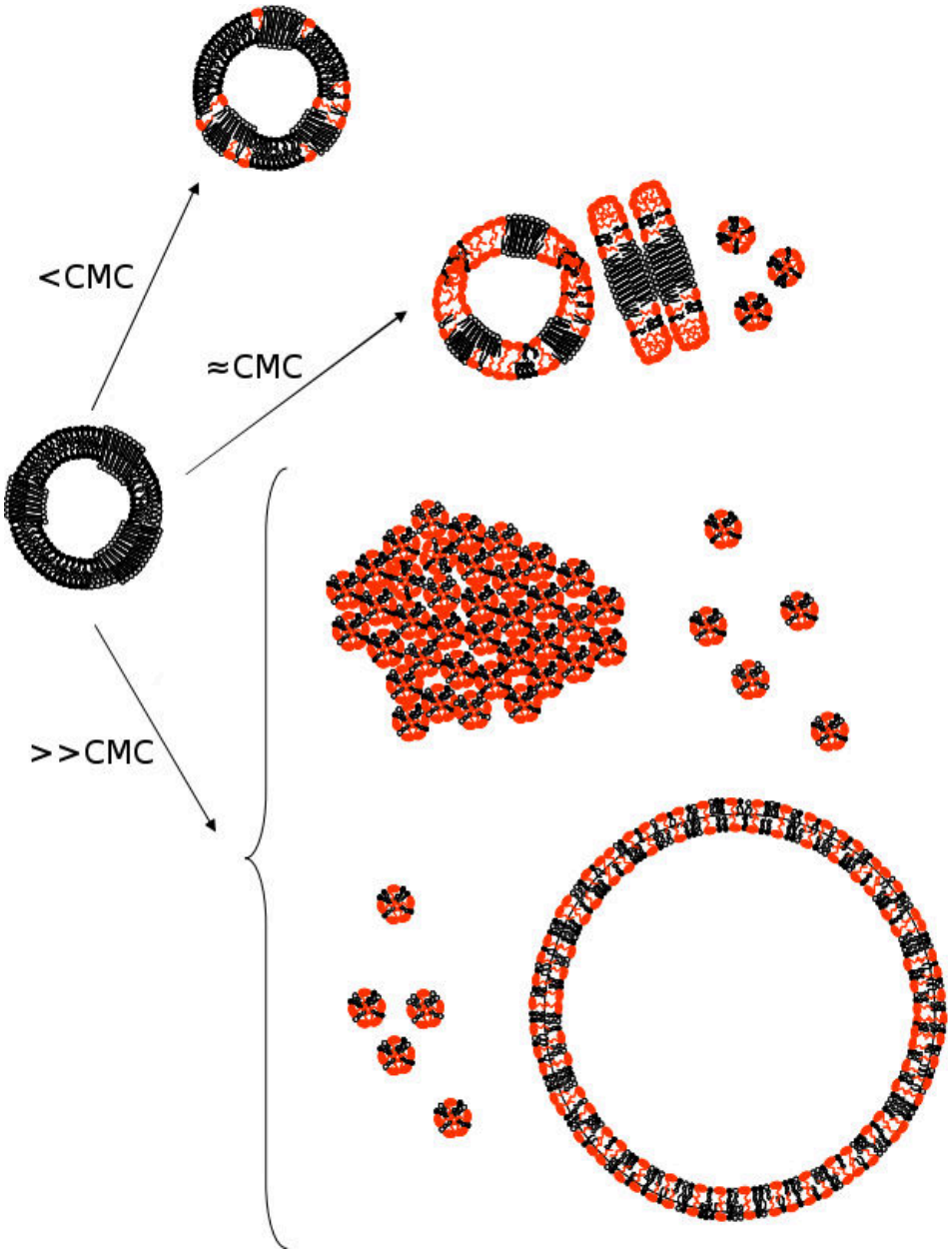


Fig. 21 : Action séquentielle de la surfactine sur une bicouche lipidique mixte.

En noir les lipides, en rouge la surfactine. Sous la CMC, la surfactine s'insère préférentiellement à l'interface entre phases fluides et plus rigides. Sa capacité de flip-flop est limitée. À concentration approchant la CMC, elle solubilise les phase fluides. Au-dessus de la CMC, elle solubilise les deux phases. Figure adaptée de Deleu *et al.* (2013).

lipopeptides pourraient également agir comme le font de nombreux antibiotiques utilisés à dose sub-efficace, en induisant, notamment, la formation de biofilm (Andersson and Hughes, 2014; Fajardo and Martínez, 2008; López *et al.*, 2009c). Enfin, la surfactine, comme déjà introduit plus tôt, serait capable d'induire la formation de biofilm à travers un rôle de signalisation spécifique *via* Spo0A (Fig. 22) (López *et al.*, 2009b).

La surfactine a été le premier lipopeptide caractérisé en tant que tel (Arima *et al.*, 1968). Il est devenu depuis le lipopeptide le mieux étudié. Il a été montré que la surfactine s'insère dans des modèles de bicouches lipidiques mixtes en plusieurs étapes à mesure que sa concentration augmente (Fig. 21) (Deleu *et al.*, 2013). À faible concentration, elle s'insère d'abord à l'interface entre phases plus fluides et plus rigides, produisant une chute de fluidité similaire à celle induite par le cholestérol. Sa capacité de flip-flop (translocation dans le feuillet lipidique opposé) est limitée. À concentration approchant sa concentration micellaire critique (CMC), elle solubilise la phase fluide. Puis au-dessus de sa CMC, elle solubilise la phase plus rigide. Cette capacité de déstabilisation des membranes à faible concentration lui permettrait de générer un efflux de potassium qui déclencherait l'activation du phosphorelais par KinC, entraînant la formation de biofilm (Fig. 22) (López *et al.*, 2009b). La surfactine est produite par une sous-population de cellules qui s'immuniserait contre son action, et interagirait donc avec la membrane des cellules ne la produisant pas, différenciant ainsi la population en deux. Ce dernier mécanisme est intéressant car il place la surfactine à l'interface entre deux régulateurs importants de *B. subtilis* : ComA et Spo0A. ComA est un régulateur permettant à *B. subtilis* d'entrer en compétence (Magnuson *et al.*, 1994). Il est activé par la kinase membranaire ComP en réponse au peptide de signalisation ComX, et fait donc partie des régulateurs de *quorum sensing*. Des études récentes décrivent deux autres rôles de la surfactine *in vivo* : elle est capable de solubiliser les vésicules émises par les bactéries à Gram positif (Brown *et al.*, 2014), et est également inductrice de la résistance systémique des plantes (Cawoy *et al.*, 2013; Khong *et al.*, 2012, 2013).

La fengycine a été découverte indépendamment par deux équipes, l'une travaillant sur *B. subtilis*, l'autre sur *B. cereus* (Nishikiori *et al.*, 1986; Vanittanakom *et al.*, 1986). Chez *B. subtilis*, la molécule a été nommée fengycine, et caractérisée comme antifongique. Chez *B. cereus*, elle a été baptisée plipastatine et caractérisée comme inhibiteur de la phospholipase A2. Comme pour la surfactine, la capacité à induire la résistance systémique des plantes de la fengycine a été depuis caractérisée (Khong *et al.*, 2012, 2013). Une étude récente montre que

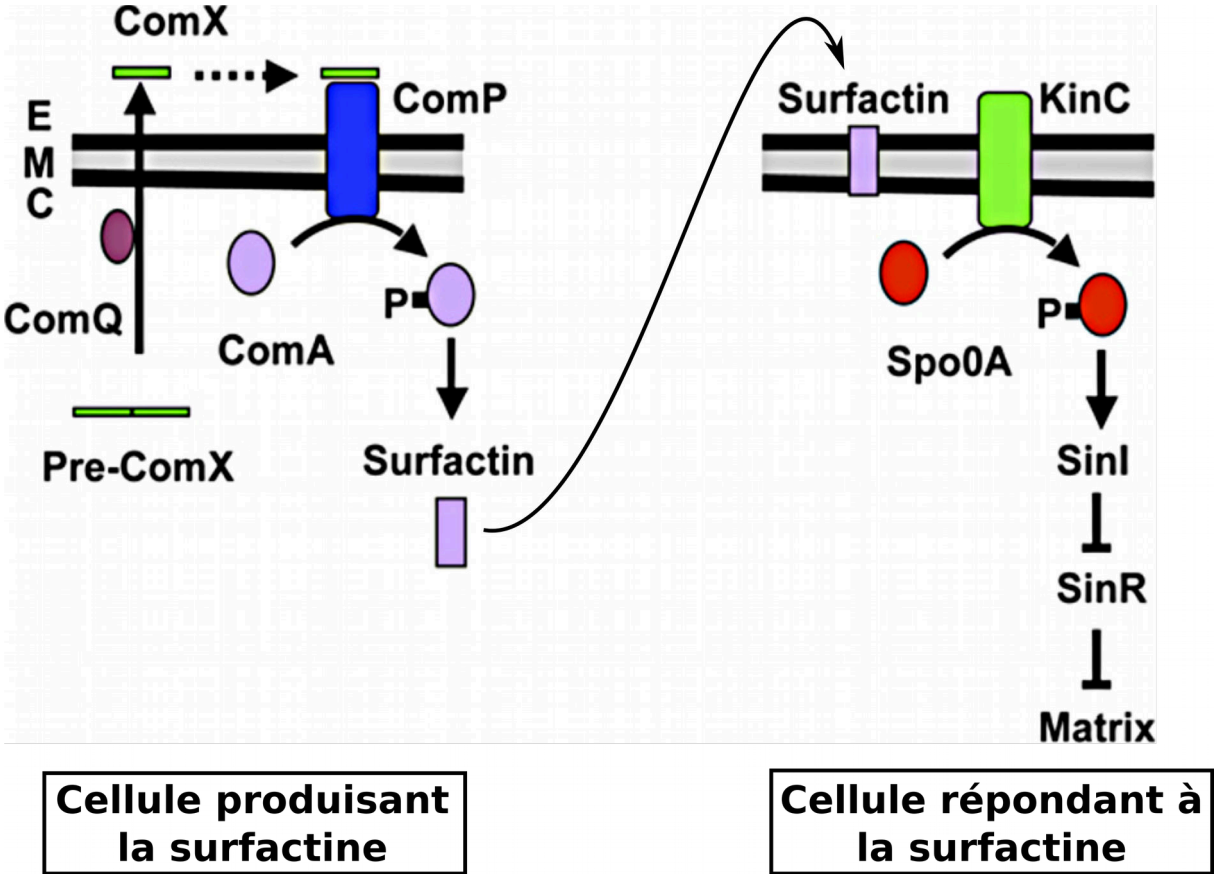


Fig. 22 : Modèle de la signalisation par la surfactine.

La surfactine est produite dans une sous-population en réponse au système de quorum sensing ComX–ComP–ComA, et active la voie Spo0A *via* KinC dans une autre sous-population. Figure daptée de López *et al.* (2009c).

la fengycine produite par une souche de *B. subtilis* commensale du nématode *Caenorhabditis elegans* protège le ver contre l'infection par d'autres bactéries à Gram positif (Iatsenko *et al.*, 2014).

La mycosubtiline a été la première molécule de la famille des iturines à être isolée, un an avant l'iturine proprement dite (Delcambe, 1950; Walton and Woodruff, 1949). Cependant, il faut attendre une trentaine d'années avant qu'elles soient caractérisées toutes deux comme lipopeptides (Peypoux *et al.*, 1976, 1978). La mycosubtiline doit son nom à son action antifongique forte, caractérisée dès sa découverte. Des études sur son insertion dans des modèles de membranes contenant de l'ergostérol, stérol majoritaire des champignons, tendent à expliquer cette capacité (Nasir and Besson, 2012). Elle a depuis également été montrée comme capable d'induire la résistance systémique de plantes (Khong *et al.*, 2012, 2013).

2 Découverte de la kurstakine

Le kurstakine est un lipopeptide produit par *B. thuringiensis*, qui a été découvert en 2000 à la surface des spores de la souche *B. thuringiensis* serovar *kurstaki* strain HD-1 (Hathout *et al.*, 2000).

Il s'agit d'un lipo-heptapeptide cyclique ou linéaire composé des acides aminés D-Thr/Gly/Ala/Ser/His/D-Gln/Gln, la thréonine 1 étant liée à un acide gras par une liaison amide (Fig. 20). Lorsqu'elle est cyclique, la sérine 4 et la glutamine 7 de la kurstakine forment une liaison lactone. La longueur et la nature de l'acide gras sont variables : iso-C_{11}, C_{12}, C_{12} β-hydroxylé, iso-C_{12}, iso-C_{13}, iso-C_{13} β-hydroxylé et C_{14} ont été détectés dans des molécules de kurstakine cycliques (Bumpus *et al.*, 2009; Hathout *et al.*, 2000). Notons que l'énantiomérie des acides aminés n'a pas été déterminée expérimentalement mais est prédite par bioinformatique (Abderrahmani *et al.*, 2011).

3 La voie de synthèse non-ribosomique

Les lipopeptides, dont la kurstakine, ne sont pas le produit d'un gène, mais sont synthétisés par de gros complexes multi-enzymatiques appelés « synthétases de peptides non-ribosomiques » (NRPS pour l'acronyme anglais).

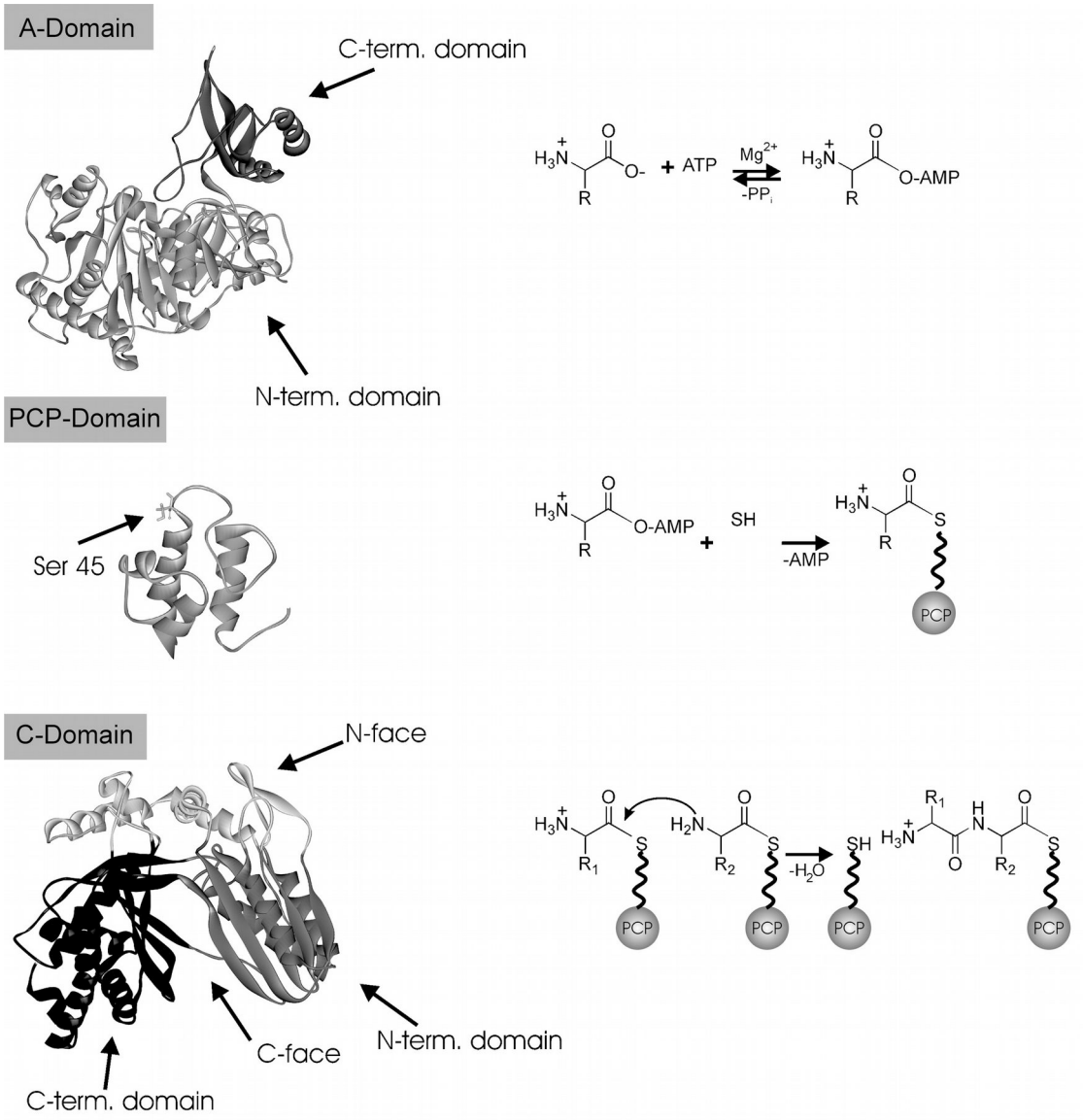


Fig. 23 : Les domaines A, PCP et C des modules NRPS.

À gauche, la structure tertiaire des différents domaines avec leurs sous-domaines. À droite, la réaction catalysée par ces différents domaines. (Finking and Marahiel, 2004).

Qu'elle soit linéaire ou itérative, la synthèse non-ribosomique de peptides se fait selon un schéma commun (Fig. 23) (Finking and Marahiel, 2004). Chaque synthétase porte un ou des module(s), chaque module étant composé de domaines catalytiques. Un module comprend ainsi un domaine A (pour « **a**dénylation »), un domaine PCP (pour « **p**eptidyl **c**arrier **p**rotein ») et un domaine C (pour « **c**ondensation »). Le dernier module contient souvent en plus un domaine Te (pour « **t**hio**e**stérase »). Le domaine A est celui qui permet à chaque module d'être spécifique d'un acide aminé. Il prend un acide aminé et une molécule d'ATP comme substrat et adényle l'acide aminé, avec pour sous-produit un ion PPi. L'acide aminé adénylé, également dit « activé », réagit ensuite avec le groupement thiol du domaine *holo*-PCP (forme liée à une molécule de 4'-phosphopantéthéine, ou PPant) pour se retrouver lié par une liaison thioester à ce domaine (d'où le nom « carrier ») (Stachelhaus *et al.*, 1996). Le domaine C catalyse alors la formation d'une liaison peptidique *via* l'attaque nucléophile de l'acide aminé lié au domaine PCP aval sur le groupement thioester de l'acide aminé lié au domaine PCP amont. De cette façon, le domaine PCP aval se retrouve chargé d'un acide aminé supplémentaire. En répétant ces trois opérations successives, la chaîne peptidique s'allonge, jusqu'à ce que le dernier domaine PCP se retrouve chargé de la chaîne complète. À ce moment-là, le domaine Te catalyse une réaction en deux étapes qui libère le produit final de la NRPS, sous forme cyclique ou linéaire selon les cas.

On note que pour que la NRPS puisse faire ses cycles, une enzyme supplémentaire lui est nécessaire : la 4'-phosphopantéthéinyl transférase (PPTase). Celle-ci se charge de lier la PPant aux domaines *apo*-PCP, les transformant en domaines *holo*-PCP. La PPant fournit le groupement thiol indispensable à l'action du domaine *holo*-PCP.

En plus de ces étapes communes de la synthèse non-ribosomique de peptides, certains domaines et/ou enzymes accessoires peuvent intervenir. Le domaine E (pour « épimérisation ») des NRPS permet notamment de changer l'énantiomérie d'un acide aminé avant son incorporation dans la chaîne peptidique. Les thioestérases de type II (TeII) permettent, quant à elles, de régénérer des domaines *holo*-PCP mal chargés. En effet, une PPTase peut lier une forme inadéquate de 4'-PPant (acyl-4'-PPant) à un domaine *apo*-PCP, le transformant en domaine *holo*-PCP, mais néanmoins inactif, ce qui a pour effet de bloquer la synthèse du produit. La TeII peut résoudre ce problème en hydrolysant le groupement acyle superfétatoire, restaurant donc un domaine *holo*-PCP actif (Schwarzer *et al.*, 2002).

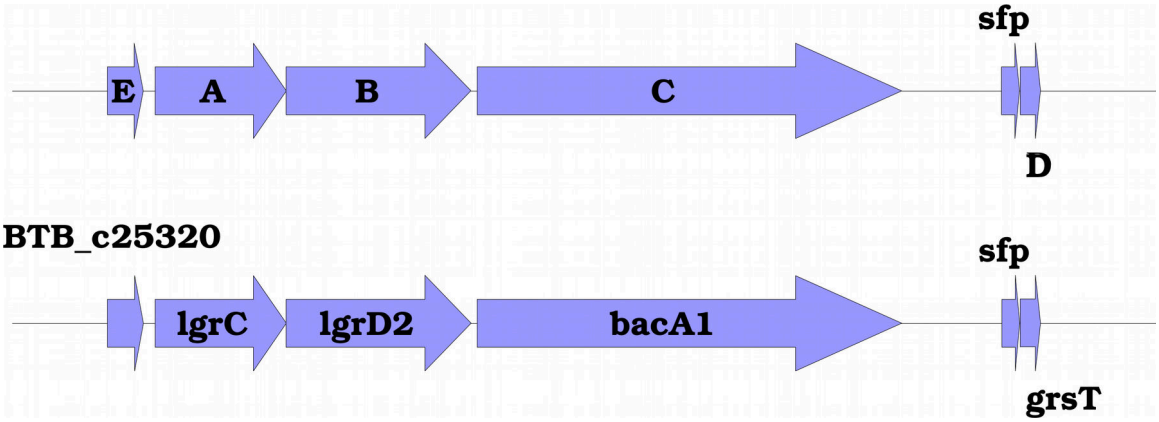


Fig. 24 : Le locus *krs* de *Bacillus thuringiensis* 407.

En haut, l'annotation du locus proposée par le laboratoire : *krsE*, *krsA*, *krsB*, *krsC*, *sfp* et *krsD*. En bas, l'annotation présente sur la séquence du chromosome déposée (numéro d'accession INSDC CP003889.1) (Sheppard *et al.*, 2013).

4 Les NRPS de *Bacillus thuringiensis* 407

Des outils bioinformatiques ont été développés spécifiquement pour les NRPS. Certains pour détecter, dans des génomes annotés, la présence de « loci NRPS » (loci dont les gènes codent pour des NRPS), d'autres pour tenter de prédire la séquence du peptide produit par ces NRPS. antiSMASH est le fruit d'une collaboration entre plusieurs laboratoires et est l'outil le plus complet disponible à ce jour : il permet de partir d'un génome annoté pour arriver à une prédiction des loci impliqués dans la production de métabolites secondaires (dont les NRPS) et de leurs produits (Medema *et al.*, 2011; Weber *et al.*, 2015).

L'analyse du génome de *B. thuringiensis* 407 par antiSMASH prédit trois loci NRPS, tous chromosomiques. Le premier code pour une NRPS impliquée dans la production de la bacillibactine, un sidérophore analogue à l'entérobactine des bactéries à Gram négatif (May *et al.*, 2000; Segond *et al.*, 2014). Le second code pour une NRPS à deux modules dont le produit semble difficile à prédire. Le troisième est le locus *krs*, détaillé à la prochaine section.

Le produit de NRPS le plus connu du groupe *Bacillus cereus* est certainement le céreulide (Agata *et al.*, 1994; Ehling-Schulz *et al.*, 2006). Ce dodécadepsipeptide est une des toxines de *B. cereus*, et est responsable du syndrome émétique. Le locus *ces*, codant pour les NRPS, est porté par un plasmide de 200 kb spécifique des souches émétiques, et donc absent de *B. thuringiensis* 407.

5 Le locus *krs*

Le locus *krs* a été lié à la production de kurstakine neuf ans après l'isolation de la kurstakine (Bumpus *et al.*, 2009; Hathout *et al.*, 2000). Le locus *krs* de *B. thuringiensis* 407 comprend 6 gènes et s'étend sur plus de 30 kb (Fig. 24). En amont du premier gène, on trouve plus d'1 kb sans annotation, puis un gène orienté de la même façon que le locus *krs* et dont le produit est prédit pour être une lipase. En aval du locus, à 177 nt de la fin du dernier gène, en orientation opposée à celle du locus *krs*, est annoté un gène dont le produit est prédit comme protéine impliquée dans l'efflux d'arsenic.

Le premier gène du locus *krs*, *krsE*, est annoté comme codant un pore membranaire impliqué dans l'efflux de macrolides. Cependant, une modélisation de structure tertiaire de KrsE par homologie, avec le programme Phyre2 (Kelley *et al.*, 2015), nous a permis de le situer proche de pores membranaires de la famille POT (pour « proton-dependent

oligopeptide transporter »), dont les membres sont impliqués dans le transport d'oligopeptides (Solcan *et al.*, 2012). De plus, il a récemment été montré que KrsE facilite l'efflux de surfactine chez *B. subtilis* (Li *et al.*, 2015). Ces données, appuyées par la présence du gène *krsE* dans le locus *krs* nous font supposer que KrsE serait en fait un pore impliqué dans l'efflux de kurstakine.

Les trois gènes suivants, *krsA*, *krsB* et *krsC*, codent pour les trois synthétases de la NRPS. KrsA porte le module responsable de l'incorporation de la thréonine, KrsB les modules pour la glycine et l'alanine, et KrsC les modules pour la sérine, l'histidine et les deux glutamines. Les modules 1 (sur KrsA) et 6 (sur KrsC) possèdent un domaine E, raison pour laquelle on suppose que la thréonine 1 et la glutamine 6 seraient sous forme D-.

Le cinquième gène du locus, *sfp*, code pour la PPTase qui fournit la PPant aux synthétases (comme détaillé précédemment).

Le sixième et dernier gène du locus, *krsD*, code pour une TeII potentiellement impliquée dans la résolution d'intermédiaires inactifs des synthétases (également détaillé précédemment).

Le locus *krs* permet ainsi la production de kurstakine, dont la fonction physiologique reste largement inconnue, seule la première étude de 2000 (Hathout et al) décrivant un possible rôle antifongique de la molécule.

CONTEXTE DE L'ÉTUDE

Les travaux menés au cours de cette thèse ont poursuivi deux objectifs. Premièrement, étudier la transcription du locus *krs* dans la souche modèle *B. thuringiensis* 407 et déterminer le(s) rôle(s) du produit de ce locus, la kurstakine. Deuxièmement, établir le régulon Spo0A de cette même souche modèle, et si possible isoler un ou des gènes de ce régulon impliqué(s) dans le cycle infectieux de *B. thuringiensis* dans son hôte *G. mellonella*. Ces deux objectifs convergent vers une tentative de mieux comprendre le comportement de *B. thuringiensis* au cours de sa phase stationnaire de croissance, *in vitro* en milieu synthétique de laboratoire, comme dans son hôte naturel, la larve d'insecte.

Comme présenté précédemment, la kurstakine est un lipopeptide produit par certaines souches du groupe *B. cereus*. Il a déjà été proposé que le locus *krs*, responsable de la synthèse de la kurstakine, soit un marqueur du groupe *B. cereus* (Béchet *et al.*, 2012). Nous avons tenté de confirmer ces données et de les préciser par des approches *in silico* fines en étudiant la répartition du locus *krs* au sein de tous les génomes séquencés, puis plus particulièrement ceux du groupe *B. cereus*. Il a été montré récemment que la transcription du locus *krs* dépend de la présence du régulateur de *quorum sensing* NprR (Dubois *et al.*, 2012), et semble modulée par SinR dans un milieu favorisant la formation de biofilm (Fagerlund *et al.*, 2014). Cependant, aucune étude à ce jour ne se consacre réellement à la régulation du locus *krs*. Nous avons donc entrepris de définir l'organisation génétique de ce locus et d'étudier précisément l'expression du promoteur en amont de son premier gène. Plusieurs phénotypes avaient été attribués à la délétion du locus *krs* (Dubois *et al.*, 2012), données que nous avons pu démentir (Annexe 1). Nous avons donc ensuite tenté de redéfinir les rôles de la kurstakine dans son contexte naturel. En collaboration avec l'équipe ProBioGEM (Polytech-Lille, université Lille I), nous avons également entrepris la caractérisation des applications possibles de la kurstakine, notamment comme molécule antifongique.

À l'inverse de la kurstakine, la littérature abonde de données sur Spo0A, régulateur majeur de la différentiation. Cependant, la quasi-totalité des données publiées concerne la bactérie modèle *B. subtilis*. De nombreux parallèles sont régulièrement faits entre *B. subtilis* et *B. thuringiensis*, reposant seulement rarement sur des données expérimentales. Or ces deux bactéries, bien que relativement proches phylogénétiquement, n'en sont pas moins différentes par de nombreux aspects. L'écologie particulière de *B. thuringiensis* (son cycle infectieux dans la larve d'insecte) ainsi que son génome, plus grand que celui de *B. subtilis* (6,5 Mb contre 4,2 Mb) et riche en plasmides, laissent supposer que Spo0A pourrait ne pas revêtir un

rôle complètement redondant entre ces deux espèces. De plus, même s’il est régulièrement supposé que Spo0A n’intervient qu’en toute fin de cycle infectieux chez *B. thuringiensis*, pour permettre la sporulation et donc la dissémination vers d’autres hôtes, les données acquises chez *B. subtilis* démontrant l’implication de Spo0A dans d’autres modes de développement laissent supposer que, finalement, Spo0A pourrait être impliqué à d’autres stades du cycle infectieux de *B. thuringiensis*. Nous avons donc d’abord tenté de montrer l’implication de Spo0A dans les autres étapes du cycle infectieux. Nous avons ensuite entrepris d’établir le régulon Spo0A de *B. thuringiensis* par trois approches : *in silico*, par prédiction des boîtes Spo0A dans le génome, *in vitro*, dans une culture pure de *B. thuringiensis* en milieu LB, et *in vivo* après infection et mort d’une larve d’insecte hôte. Dans notre analyse de ces régulons, nous souhaitions nous pencher sur les spécificités du régulon de *B. thuringiensis* par rapport à celui de *B. subtilis*, et particulièrement sur les gènes du régulon que nous aurions défini comme potentiellement impliqués dans le cycle infectieux de *B. thuringiensis* dans son hôte.

RÉSULTATS

Chapitre I - Transcription et rôle de la kurstakine

La kurstakine est un lipo-heptapeptide cyclique synthétisé de façon non-ribosomique par *B. thuringiensis* (Abderrahmani *et al.*, 2011; Bumpus *et al.*, 2009; Hathout *et al.*, 2000). Avant le travail présenté ici, huit articles de recherche indexés (d'après la base de données PubMed, le 01/10/2015) citaient la kurstakine, dont seulement deux réellement centrés sur la kurstakine, sous un angle majoritairement biotechnologique (Abderrahmani *et al.*, 2011; Hathout *et al.*, 2000). Une seule donnée sur la transcription du locus *krs* était publiée, montrant que celle-ci nécessitait la présence du régulateur de *quorum sensing* NprR (Dubois *et al.*, 2012). Dans ce même article, plusieurs phénotypes avaient été attribués à la délétion du locus *krs*, données que nous avons par la suite démenties (Annexe 1). Nous avons donc choisi d'aborder la kurstakine dans son contexte naturel, en étudiant à la fois la régulation transcriptionnelle du locus *krs* et les rôles potentiels de son produit, la kurstakine, dans le développement de *B. thuringiensis*.

1 Article 1 : « Transcription and role of kurstakin, a lipopeptide produced by *Bacillus thuringiensis* »

1.1 Phylogénie du locus *krs*

L'utilisation du logiciel BLASTats, présenté au chapitre III des résultats, nous a permis de rechercher exhaustivement les organismes séquencés dont le génome porte un locus *krs* supposé fonctionnel (de tels génomes seront désignés *krs*$^+$ par la suite). Le locus *krs* n'a été retrouvé qu'au sein du groupe *Bacillus cereus*. Cependant, au sein du groupe, sa répartition n'est pas homogène : 45 % des génomes de *B. thuringiensis* sont *krs*$^+$, contre 12 % pour *B. cereus* et 0 % pour *B. anthracis* (Tab. 1). Le faible nombre de génomes séquencés dans les autres espèces du groupe ne permet pas de tirer d'information statistique, néanmoins on note que le locus *krs* n'a pas été retrouvé dans les uniques génomes disponibles de *B. cytotoxicus* ni de *B. toyonensis*. En revanche, un des deux génomes de *B. mycoides* est *krs*$^+$, ainsi qu'un des deux génomes de *B. weihenstephanensis*. Dans ces deux derniers cas, le locus est présent sur de gros plasmides, pBMX_1 (360 kb) et pBWB401 (400 kb) respectivement, flanqués d'éléments d'insertion ou de gènes viraux semblant marquer une mobilité du locus *krs* médiée par des transposons ou des phages.

La construction d'un arbre phylogénétique des souches possédant un génome *krs*$^{+}$ indique qu'elles se trouvent exclusivement dans le second sous-groupe du groupe *Bacillus cereus*, à l'exception des deux souches ayant un locus plasmidique qui appartiennent au troisième sous-groupe (Fig. 1) (Tourasse and Kolstø, 2008).

1.2 La régulation transcriptionnelle de l'opéron *krs* est complexe

Après avoir déterminé l'organisation du locus *krs* en opéron (Fig. 2), nous nous sommes intéressés à sa régulation transcriptionnelle.

Nous avons dans un premier temps déterminé la position d'un démarrage de transcription de *krs* en amont du premier gène, 327 nt en amont de son codon start (Fig. 3A). La dissection du promoteur *krs* sur plasmide nous a ensuite permis de définir la zone promotrice de *krs* et de tester la fonctionnalité de certaines séquences repérées par bioinformatique : de 5' vers 3', deux boîtes SinR potentielles, un palindrome composé de répétitions inverses exactes de 15 nt espacées par 6 nt, et une potentielle boîte Spo0A en aval du démarrage de transcription (Fig. 3B, C et D). Les deux boîtes SinR semblent ne pas intervenir dans l'expression de *krs*, car leur délétion est restée sans effet. Nous avons montré qu'en revanche, la boîte Spo0A réprime l'expression de *krs*, en accord avec les cinétiques d'expression de *krs* dans un mutant de délétion de *spo0A* (Fig. 3E). Un élément clé, encore mal compris, de la transcription de *krs* est le palindrome. Lorsque la répétition sens et la région 5' amont sont délétées, la transcription de *krs* augmente. En revanche, lorsque cette même répétition est délétée en gardant la région amont, *krs* n'est presque plus transcrit. Nous avons finalement défini la zone promotrice minimale, qui remonte 163 nt en amont du démarrage de transcription. La délétion de 86 nt supplémentaires en 5' fait chuter la transcription de moitié, et la délétion de 32 nt supplémentaires ne permet plus la transcription du tout.

Nous avons pu montrer que la transcription de l'opéron est fortement dépendante des conditions de croissance. En milieu liquide agité, nous avons mesuré, à l'échelle de la population, des différences de timing d'expression, notamment entre milieux inducteurs de différenciation (LBP et HCT) et milieux non-inducteurs (LB et BHI) (Fig. 4A). Nous avons également constaté, à l'échelle cellulaire cette fois, que si l'expression de *krs* est homogène dans la population en milieu liquide agité, elle est bimodale dans le biofilm cultivé sur milieu

solide (Fig. 4B).

1.3 La kurstakine est impliquée dans la structuration du biofilm

Nous avons montré que le produit du locus *krs*, la kurstakine, est impliqué dans la structuration correcte du biofilm de *B. thuringiensis*. Le biofilm formé à l'interface air–liquide par un mutant de délétion ne produisant pas de kurstakine, en tube de verre à 48 heures, est moins dense que le biofilm formé par la souche parentale (Fig. 5A). Dans d'autres conditions, en plaque PVC à 24 heures, la densité du biofilm n'est pas affectée mais la structure des pellicules diffère grandement (Fig. 5B). En plus de cette différence d'apparence, l'application d'une pression perpendiculaire au biofilm révèle des réactions différentes entre les deux souches : le biofilm formé par la souche parentale est traversé aisément (la pellicule se rompt au point de contact) alors que le biofilm formé par la souche mutante reste intègre et est emporté au fond par la pression (Film S1). La culture de biofilms sur filtre polycarbonate nous a permis d'observer en microscopie confocale que, dans ce modèle également, la souche mutante et la souche parentale forment des biofilms d'épaisseur et de structures très différentes (Fig. 5C).

Article 1 : « Transcription and role of kurstakin, a lipopeptide produced by *Bacillus thuringiensis* »

Manuscrit soumis à Research in Microbiology

Genetic and functional analyses of *krs*, a locus encoding kurstakin, a lipopeptide produced by *Bacillus thuringiensis*.

Sébastien Gélis-Jeanvoine[a], Alexis Canette[a], Michel Gohar[a], Thibault Caradec[b], Christelle Lemy[a], Myriam Gominet[c], Philippe Jacques[b], Leyla Slamti[a]*, Didier Lereclus[a]*

[a]Micalis Institute, INRA, AgroParisTech, Université Paris-Saclay, 78350 Jouy-en-Josas, France

[b]University of Lille, EA 7394, ICV-Institut Charles Viollette, ProBioGEM team, Polytech'Lille, avenue Langevin, 59655 Villeneuve d'Ascq, France

[c]Institut Pasteur, CNRS URA 2172, Unité de Biologie des Bactéries Pathogènes à Gram positif, 75015 Paris, France

sebastien@gelis.ch, michel.gohar@jouy.inra.fr, myriam.gominet@pasteur.fr, leyla.slamti@jouy.inra.fr, Didier.Lereclus@jouy.inra.fr

*Correspondence and reprints

Abstract

Bacteria of the *Bacillus* genus are able to synthesize several families of lipopeptides. These small molecules are the product of non-ribosomal peptide synthetases. In 2000, it was found that *Bacillus thuringiensis*, an entomopathogenic bacterium of the *Bacillus cereus* group, produced a previously unknown lipopeptide: kurstakin. Genomic analyses reveal that the *krs* locus, encoding the kurstakin synthetases, is specific to the *B. cereus* group but is unevenly distributed within this group. Previous work showed that *krs* transcription requires the necrotrophism quorum-sensor NprR. Here, we demonstrated that the genes of the *krs* locus form an operon and we defined its transcription start site. Following *krs* transcription at the population and single-cell levels in multiple culture conditions, we depicted a condition-dependent transcription pattern, indicating that production of kurstakin is subject to environmental regulation. Consistent with this idea, we found *krs* transcription to be regulated by another master regulator, Spo0A, suggesting that *krs* expression is fine-tuned by integrating multiple signals. We also reported an unknown DNA palindrome in the *krs* promoter region that modulates *krs* expression. Due to their surfactant properties, lipopeptides could play several physiological roles. We showed that the *krs* locus was required for proper biofilm structuration.

Keywords: Biofilm; Transcriptional regulation; Spo0A; *Bacillus cereus* group

Introduction

Lipopeptides are well-known secondary metabolites produced by bacteria of the *Bacillus* and *Pseudomonas* species. They are the products of big enzymatic complexes, the so-called non-ribosomal peptide synthetases (NRPSs). Lipopeptides produced by *Bacillus* belong to four structurally diverse families: surfactins, fengycins, iturins and kurstakins [1]. These molecules present physical properties of interest for various industrial sectors and are used as biosurfactants, emulsifiers, antifungals, antibacterials, antivirals, antitumorals, immunosuppressants and plant elicitors [2–4]. This has led to extensive studies about lipopeptide properties, but little is known about their biological role and the regulation of their production (see [5] for a review). Most data currently available show that transcription of lipopeptide synthetases is often regulated by two-component and/or quorum-sensing systems [5]. Transcription of *srfA*, the locus responsible for the synthesis of the most studied lipopeptide, surfactin, is regulated by the ComXPA quorum-sensing system in *Bacillus subtilis* [6,7]. Interestingly, despite being quorum-regulated, *srfA* was shown to be transcribed in only a small part of the cells in the population [8]. Surfactin is then perceived as a signal by the non-producer bacteria, which increases phosphorylation of Spo0A *via* the kinase KinC. Spo0A~P, the active form of the master regulator of differentiation in *B. subtilis*, will then trigger biofilm formation, cannibalism and sporulation [9].

Bacillus thuringiensis is an entomopathogenic bacterium from the *Bacillus cereus* group [10]. *B. thuringiensis* kurstaki HD-1 was shown to produce kurstakins, molecules that constitute a new class of *Bacilli* lipopeptides [11]. Our model strain *B. thuringiensis* 407 also harbors a *krs* locus and produces kurstakin [12,13]. The *krs* locus (Fig. S1) encodes the NRPS responsible for kurstakin synthesis and contains 6 genes: *krsE*, predicted to encode a permease potentially involved in the efflux of kurstakin [14,15]; the 3 kurstakin synthetases *krsA*, *krsB* and *krsC; sfp*, a phosphopantheteine transferase; and *krsD*, a type II thioesterase. It was shown that *krs* transcription is activated at the onset of sporulation by the quorum sensor NprR, a major regulator required for the necrotrophic stage during the life cycle of *B. thuringiensis* (*i.e*. its capacity to survive in the

cadaver of an insect host) [13]. In contrast, SinR was found to negatively affect *krs* transcription in a biofilm-promoting medium [16].

Lipopeptides were shown to be involved in various bacterial processes that often appear to be linked to the adaptation of a bacterium to a particular niche (see [5] for a review). In the case of *B. thuringiensis*, kurstakin could be involved in necrotrophism by increasing the spreading ability of the population, by promoting biofilm formation, or by protecting *B. thuringiensis* against host innate immunity or the host microbiota. The two latter proposed functions of kurstakin rely on its putative ability to destabilize biological membranes of both prokaryotes and eukaryotes, as it is the case for all other known lipopeptides (reviewed in [5]). It was notably shown that two lipopeptides induced the lysis of amoeba, protozoans that strongly resemble phagocytes, the main actors of eukaryotes innate immunity [17].

The potential involvement of kurstakin in the life cycle of *B. thuringiensis* and its potential industrial interest prompted us to investigate its regulation and roles. Here, we demonstrated that the *krs* locus is an operon that is unevenly distributed within the *B. cereus* group. We showed that, in addition to being activated by NprR, the transcription of *krs* is affected by Spo0A and a sequence in the upstream intergenic region of *krs* encompassing a DNA palindrome. Finally, we bring evidence that kurstakin is involved in proper biofilm structuration.

Materials and methods

Bacterial strains and growth conditions

Bacillus thuringiensis str. 407 Cry$^-$ (referred to as *B. thuringiensis* 407 in the text) is an acrystalliferous strain cured of its *cry* plasmid [18]. Strains DH5α and TG1 of *Escherichia coli* were used as hosts in the process of plasmid construction. Plasmid DNA was then prepared from the *dam dcm E. coli* strain ET12567 (Stratagene, La Jolla, CA, USA) and transformed by electroporation of *B. thuringiensis* 407 as described previously [18]. *B. thuringiensis* 407 derivative strains used in this study are listed in Table S1. *E. coli* strains were grown in shaken LB at 37°C [19]. *B. thuringiensis* 407 was grown in LB, BHI (adapted from [20,21]), LBP or HCT at 30°C. LBP and HCT are known to promote *B. thuringiensis* differentiation [22,23]. When required, the medium was supplemented with antibiotics at the following concentrations: 100 µg/mL ampicillin (for *E. coli*), 200 µg/mL kanamycin, 200 µg/mL spectinomycin, 10 µg/mL erythromycin (for *B. thuringiensis* 407).

DNA manipulations

Plasmid DNA was extracted from *E. coli* by standard alkaline lysis method and followed by standard purification on a silica column, using the PureYield Plasmid Miniprep System (Promega, Madison, WI, USA). Restriction enzymes and T4 DNA ligase (New England Biolabs, USA) were used according to the manufacturer's instructions. Oligonucleotide primers (Table S3) were synthesized by Sigma-Proligo (Paris, France). PCRs were performed in thin-walled tubes in a 2720 Thermak cycler (Applied Biosystem, USA). Amplified fragments were purified on silica columns using the Wizard SV Gel and PCR Clean-Up System (Promega, Madison, WI, USA). Digested fragments were separated on a 0.8% TAE-agarose gel containing ethidium bromide and purified with the same kit. Sanger sequencing of plasmid and PCR DNA was done by Beckman Coulter Genomics (Takeley, UK).

Directed plasmid mutagenesis

To construct the pHT304-18ΩP$_{krs}$(0A*)'-*lacZ* plasmid (Table S2), the pHT304-18ΩP$_{krs}$(FL)'-*lacZ* plasmid was used as the template of a PCR with the high-fidelity Phusion polymerase (NEB) and the SG74 and SG75 primers (Table S3). In this plasmid, the putative Spo0A box in the *krs* promoter region was mutated from TTTGACGAA to TTTcAgGAA to prevent Spo0A binding [24]. The whole plasmid, integrating the two point mutations, was synthesized during the reaction. The template plasmid was subsequently degraded by addition of DpnI, a restriction enzyme that recognizes methylated GATC motifs. The PCR product was transformed in *E. coli* TG1 and the recombinant clones were selected for resistance to ampicillin. The presence of the mutations was verified by Sanger sequencing.

Construction of the B. thuringiensis 407 recombinant strain

The *krsC* gene, encoding KrsC, the largest of the three kurstakin synthetases, was disrupted by insertion of the promoter-less and terminator-less *lacZ* gene, by homologous recombination (as described previously [25]), using the thermosensitive plasmid pRN5101Ω*krsC*::*lacZ* (Table S2). This strain was verified by PCR and by Sanger sequencing of the regions used for recombination. This strain was used both to monitor *krs* transcription from the chromosomal promoter and as a Δ*krsC* deletion mutant. This mutation most likely has a polar effect on the two downstream genes *sfp* and *krsD*. These genes are predicted to encode a phosphopantheteine transferase and a type II thioesterase respectively, two enzymes known to be involved in non-ribosomal peptide synthesis. It is therefore not surprising to find them within the *krs* locus and it is reasonable to think that the products of these genes are involved in (albeit not necessarily exclusively in) kurstakin synthesis.

RT-PCR analysis of the krs *locus*

B. thuringiensis cells were grown in shaking LB at 30°C until 5 hours after the onset of transition phase (t5) and a 2 mL sample was mixed with 2 mL of RNAlater (Life Technologies Ambion, USA) before storage at -20°C. After gentle thawing, the sample was centrifuged and the

pellet was resuspended in 1 mL of Trizol (Life Technologies Ambion, USA). Resuspended cells were disrupted by bead beating (Mini 0.1 mm zirconia/silica beads, Biospec Products) for 2 x 45 s at 6.5 M/s in a Fastprep 24 (MP Biomedicals). The supernatant was transferred to a clean tube and 100 µL of 1-bromo-3-chloropropane (Sigma) were added. The suspension was vortexed, incubated for 10 min at room temperature and centrifuged for 15 min at 12000 x g at 4°C. The aqueous phase was then transferred to a clean tube to which 0.1 volume of sodium acetate and 0.7 volume of isopropanol were added. The suspension was mixed and incubated at 4°C for 30 min to precipitate nucleic acids. Nucleic acids were then pelleted by centrifugation (20 min at 12000 x g at 4°C), washed twice with 75% ethanol, air dried and resuspended in water. Possible traces of contaminating DNA were removed by DNase treatment (TURBO DNase, Life Technologies Ambion, USA) according to the manufacturer's instructions. RNA quantity and integrity was assessed *via* A_{260}/A_{280} and A_{260}/A_{230} (NanoDrop 2000, Thermo Scientific) and RNA Integrity Number (Nano chip, BioAnalyzer 2100, Agilent Technologies) [26]. Total RNA was reverse transcribed into cDNA with AffinityScript qPCR cDNA synthesis kit (Agilent), using the provided random primers. 2 µL of the reaction were used as a template in a classic 20 µL 30 cycles PCR with Taq polymerase (NEB). Primers pairs were designed to amplify over intergenic regions, except for the internal *krsB* positive control. RT-PCR primers are listed in Table S3. Primers pairs are numbered according to Fig. 2A (*i.e.* RT1_FW and RT1_REV correspond to reaction 1 in the figure).

Primer extension

Total RNA for primer extension assay was prepared from a shaken HCT culture at 37°C at t2 as described previously [27]. To detect the transcription start site of the *krs* locus, primer extension was performed using the oligonucleotide tss_bc2450 reverse complementary to the 5’ end of the *krsE* gene (Table S2). The same oligonucleotide was used to prime dideoxy sequencing reactions from the *B. thuringiensis* 407 chromosome.

β-Galactosidase assay

Cells were cultured in shaking LB at 30°C and β-galactosidase activity was assayed as described previously [28]. Specific activities were expressed as β-galactosidase units per milligram of total protein. Means and standard deviations of at least three independent experiments are presented. When appropriate, Student's t-Test (unpaired samples, two-sided, unequal variances, confidence interval = 0.95) was used to assess the significance of observed differences. Resulting *p* values were corrected using Benjamini, Hochberg, and Yekutieli's method.

Air-liquid biofilm formation

Biofilms were allowed to form in HCT in 6 mL glass tubes and 48-well polyvinylchloride (PVC) plates (Costar 3548, Cambridge, USA) as described previously [29]. 24-hour microtiter plate incubations at 30°C allowed for best biofilm structures observation, while 48-hour glass tube incubations at 30°C allowed for formation of robust and more easily manipulated pellicles. To quantify biofilm formation, the planktonic phase was discarded with a Pasteur pipette and the total biofilm (ring and pellicle) was resuspended in 1 mL of PBS to measure the OD_{600}. To test the resistance of pellicles to pressure, we melted the tip of a Pasteur pipet until it formed a ball of 5 millimeter of diameter and used it to push on the pellicles (see Movie S1).

Biofilm formation on solid medium

Biofilms on solid medium were grown as follows (adapted from [30]). Strains were streaked on LB agar plates and incubated at 30°C overnight. The next day, a few colonies were scraped and resuspended in PBS to an OD_{600} of 1, and 2 µL of this suspension was spotted at the center of the plate. The droplet was allowed to penetrate the agar (~5 min) before incubation for 72 h at 30°C. After incubation, the colony was scraped and resuspended thoroughly in 1 mL of 4% formaldehyde in PBS. Cells were incubated ~5 min at room temperature and centrifuged for 30 seconds at

13.000 x g. The pellet was resuspended in 1 mL of PBS and centrifuged 5 min at 600 x g to wash out formaldehyde. The pellet was finally resuspended in 100 µL of GTE buffer [30].

Flow cytometry analyses

Flow cytometry analyses were adapted from [31]. Fluorescence was measured on a CyFlow Space cytometer (Partec, France). GFP fluorescence driven by the *krs* promoter was recovered after excitation by a solid blue-laser emitting at 488 nm combined to a 500-nm long pass dichroic mirror and a 527-nm band pass filter (512–542) (FL1 Channel). Measures were acquired using logarithmic gains and detector settings, adjusted on a sample of reporter-less cells, to define cellular auto-fluorescence. Gating on FSC⁄SSC was used to discriminate bacteria from background particles. For each sample, 100.000 gated events were acquired. Data were collected with the FloMax software (Partec, France) and figures were generated with the Weasel 2.0 software (WEHI, USA).

Confocal laser scanning spectral microscopy

Overnight HCT precultures (30°C) were used to inoculate fresh HCT incubated at 30°C to an OD_{600} of 0.05. When the cultures reached an OD_{600} of ~1, they were diluted to an OD_{600} of 0.01 in 10 mL fresh HCT in 30 mL vials. 0.4 µm polycarbonate isopore filters (Millipore) were punched out to an appropriate diameter, UV-sterilized and set to float in the 30 mL vials. Samples were then incubated for 48 hours at 30°C. The filters were gently transferred on 35 mm low µ-dishes (Ibidi) containing 800 µl of HCT medium supplemented with fluorescent markers. The nucleic acid dye Syto9 (Life Technologies) was used at a final concentration of 6.67 nM and the lectin stain Concanavalin A (ConA) Alexa Fluor 633-conjugated (glycoconjugates detection, Life Technologies) was used at a final concentration of 200 µg/mL. Samples were incubated in the dark at room temperature for at least 10 min. The dishes were then mounted on the motorized stage of an inverted confocal microscope (Leica TCS SP8 AOBS, Leica Microsystems) at the MIMA2 platform (http://www.jouy.inra.fr/mima2/). The dishes were scanned using a 40×/0.8 N.A. water immersion

objective lens because of its interesting working distance (3300 µm) allowing thick z-stacks at a convenient magnification level. Syto9 was excited at 488 nm using an argon laser (output power at 30%, AOTF at 0.5%) and the emitted “green” fluorescence was recorded from 498 to 600 nm with a PMT detector with a 750 V gain. Simultaneously, ConA was excited at 633 nm using a He-Ne laser (AOTF at 5%) and the emitted “red” fluorescence was recorded from 643 to 750 nm with a hybrid detector with 20% digital gain. 2D sections were acquired at 600 Hz with numerical zoom 1, 512 x 512 xy image definition and a z-step of 1 µm. Three-dimensional projections of biofilms were then reconstructed with the “easy 3D blend” module of the IMARIS 7.7.2 software (Bitplane).

Infection of Galleria mellonella *larvae*

Larvae infections were carried out as described in [13], except that the inoculum was set to 2.10^5 CFU.

Bioinformatics analyses

Nonribosomal peptide synthetases were detected in the reference genome of *B. thuringiensis* 407 [32] (chromosome: INSDC CP003889.1, plasmids: INSDC CP003890.1 to CP003898.1) with antiSMASH 3 [33,34].

The presence of all *krs* genes was investigated in all completely sequenced species with BLASTp (v. 2.2.30+) on the unfiltered nr database of NCBI (as of 2015-06-01) [35], using *B. thuringiensis* 407 sequences as queries.

Distribution of kurstakin synthetases among the *B. cereus* group was assessed with an in-house Python script (BLASTats, https://github.com/seb-ksl/BLASTats/). Briefly, this script submits a protein sequence to NCBI's tBLASTn using the nr/nt database. Hits whose best high scoring segment pair (HSP) was over 80% sequence identity and over 85% query coverage were considered significant. These thresholds allowed detection of *krs* genes with no false positive on other NRPS loci. The name of the strain(s) generating each significant hit was then parsed from hit titles. Strains

were grouped by species and compared to the number of completely sequenced genomes available for each species, giving the relative presence of query protein among selected species. Significant hits for each gene of the *krs* locus were further aligned with ClustalΩ (v. 1.2.1) [36] and alignments were concatenated. The resulting alignment was curated with Gblocks (v. 0.91b) [37,38] and an approximately-maximum-likelihood phylogenetic tree was built with FastTree (v. 2.1.7-1) [39,40]. Trees were then edited with FigTree (v. 1.4-2) (http://tree.bio.ed.ac.uk/software/figtree/).

Upstream intergenic sequences of all found *krs* loci were retrieved from Genbank using an in-house Python script and aligned with ClustalΩ (v. 1.2.1) [36]. The alignment was then edited with Jalview (v. 2.9.0b2) [41].

Results

The krs *locus is not evenly distributed within the* B. cereus *group*

Database searches of the *krs* gene products with all the available GenBank sequences did not yield any significant hit outside the *B. cereus* group, suggesting that *krs* is specific to this group. However, not all strains contain an intact *krs* locus. The *B. thuringiensis* species exhibits the highest frequency of strains containing an intact *krs* locus (Krs^+ strains), with 52% (12/23) positive strains, *vs* only 12% (3/26) for *B. cereus* and 0% (0/31) for *B. anthracis* (Table 1, Table S4). The absence of *krs* in all *B. anthracis* strains is explained by the high monomorphism of this species [42]. It should also be noted that the 23 *B. thuringiensis* genomes encompassed the *Bacillus bombysepticus* str. Wang genome [43]. Indeed, although *B. bombysepticus* was originally described as a separate species, its 16S RNA and genomic phylogenies, its ability to form parasporal inclusions as well as its entomopathogenicity towards *Bombyx mori* strongly indicates that it is most likely to be a misclassified strain of the *B. thuringiensis* species.

Among the 23 completely sequenced *B. thuringiensis* genomes, three are redundant (genome alignments point them as the same strain): Bt407, CT-43 and IS5056. It is therefore not surprising to find all three are Krs^+, and the frequency within *Bacillus thuringiensis* should be corrected to 10 out of 21. Two sequences of *B. thuringiensis* serovar *kurstaki* str. YBT-1520 are available, but surprisingly one of the two sequences (INSDC CP004858.1) seems to have a truncated non-functional *krsC*, while the other one (INSDC CP007607.1) has a Krs^+ locus. It should also be noted that *B. thuringiensis* serovar *galleriae* str. HD-29 harbors all *krs* genes, but its *krsB* gene seems disrupted, making the locus Krs^-. This event is likely recent because sequences of the other *krs* genes have not yet diverged from those of *B. thuringiensis* 407 (>90% overall identity). Although these last two strains (YBT-1520 and HD-29) were counted as Krs^-, we cannot exclude the possibility of sequencing errors.

A tree was built from all Krs^+ loci and reveals the high conservation of this locus (Fig. 1). A phylogenetic tree of the *B. cereus* group was proposed by Guinebretière *et al.* [44] and further

refined by Tourasse *et al.*, based on available MLST/AFLP/MLEE data [45]. These trees contain seven clusters. The topology of the *krs* tree is consistent with the existing trees and shows that most Krs^+ strains belong to cluster II of the *B. cereus* group. *B. weihenstephanensis* KBAB4 and *B. mycoides* strain ATCC 6462 are the only Krs^+ strains that fall in cluster III, and they are also the only two strains in which the *krs* locus is located on a plasmid. In *B. weihenstephanensis* KBAB4 the *krs* locus lies on the 400 kb plasmid pBWB401, flanked by two putative IS240 insertion elements. In *B. mycoides* ATCC 6462 the *krs* locus is found on the 360 kb plasmid pBMX_1, flanked by putative phage genes. To date, these are the only sequenced plasmids containing *krs*.

The krs *locus forms a six-gene operon*

The *krs* locus of *B. thuringiensis* 407 consists of six genes spanning over 30 kb on the chromosome (Fig. 2A). Very short (2 and 13 nt) or absence of intergenic space suggested that the first four genes were co-transcribed. To verify this hypothesis and examine if the last two genes belonged to the same transcriptional unit, we investigated the transcripts of the *krs* locus. Total RNA extracted from *B. thuringiensis* 407 cells was used to perform RT-PCR with primer pairs amplifying all intergenic sequences (Fig. 2A and 2B, Table S3). We amplified fragments of the expected size with each primer pair, indicating that consecutive genes were co-transcribed. This result, combined with recent unpublished RNA-seq data from our team and *in silico* analyses showing no detection of internal terminators, strongly suggests that *krs* is an operon. However, based on the data presented here, we cannot rule out the presence of alternative promoters and/or internal attenuators.

Analysis of the krs *operon promoter region*

We then performed primer extension experiments using primer tss_bc2450 and RNA extracted from *B. thuringiensis* 407 cells to identify transcription start sites (TSSs) upstream from *krsE*. Figure 3A shows a single TSS 327 bp upstream of the *krsE* translation start site. Putative -10

and -35 boxes resembling the TTGACA and TATAAT consensus of *B. subtilis* [46] may suggest that transcription from this promoter depends on σ^A.

In silico analysis of the intergenic region upstream of *krsE* revealed three notable features: two head-to-tail oriented potential SinR binding sequences, a potential Spo0A binding site and an exact DNA palindrome composed of two 15 nt inverted repeats spaced by 6 nt (Fig. 3B). Conservation of these features was assessed by aligning upstream intergenic sequences of all found *krs* operons in the nr/nt database (Fig. S2). Overall this region seems very well conserved in 15 of the 17 aligned sequences, the 2 plasmidic sequences clearly forming an out-group. Within the 15 chromosomal sequences, the putative -10 and -35 boxes, the TSS, the downstream potential SinR binding sequence and the putative Spo0A binding site are fully conserved. Two alleles of the upstream potential SinR binding sequence are found, both compatible with the published SinR box [47]. The palindrome is the least conserved feature, but is still present in 11 out of the 15 chromosomal sequences. Within these 11 sequences, it is interesting to note that two alleles are found, both forming an exact palindrome. Indeed, in the sequences of *B. thuringiensis* strains HD-1, HD73 and YBT-1520, G5 and its cognate C32 transitioned to A5-T32.

To assess the functionality of these features, we measured β-galactosidase activity generated by transcriptional fusions between various fragments amplified from this intergenic region and the reporter gene *lacZ* on the pHT304-18Z plasmid 3 hours after the end of exponential phase (t3) in HCT (Fig. 3C). In this medium, maximal expression of *krs* is reached at t3 (Fig. 4A), hence its use as a reference point. Our reference full-length fusion (FL) spanned from -739 to +84 relative to the TSS, covering the entire upstream intergenic region and the beginning of *krsE*. All consecutive fusions share the same 3' end at +84. Transcription of the FL fusion followed the same pattern as the chromosomal fusion, albeit at higher levels (Fig. S3). The map1 fragment, in which both putative SinR binding sequences were removed presented approximately the same activity as the full length fragment (0.83 +/- 0.06), suggesting that these sequences are not functional in the conditions used in this experiment (Fig. 3D). These results are supported by data showing that in our conditions,

P_{krs}(FL) activity is unchanged in a *B. thuringiensis* 407 Δ*sinIR* mutant strain compared to the WT strain (Fig. S4). Removing an additional 290 nt region encompassing the palindrome (fusion map2) resulted in a twofold increase in the β-galactosidase activity (1.98 +/- 0.08), indicating that this sequence negatively affects the transcription of *krs*. Removing 44 and 130 additional nucleotides (fusions map3 and map4, respectively) induced a threefold drop compared to the map2 fusion, suggesting the presence of an activating region immediately downstream of the palindrome. Removing 32 more nucleotides from the 5' end, leaving only 45 bp upstream of the TSS (fusion map5), generated an almost inactive fusion (0.03 +/- 0.01). Altogether, these results point the palindrome and its upstream region as playing a negative role on transcription from this promoter.

krs *expression is modulated by Spo0A*

The presence of a predicted Spo0A box between the TSS and the translation start codon suggested that Spo0A~P could act as a direct repressor of *krs* transcription [48]. To assess the involvement of Spo0A in *krs* transcription, we first monitored transcription of the FL fusion in a Δ*spo0A* strain between t2 and t5, a time frame within which an effect of Spo0A on *krs* transcription was most likely to occur (Fig. 3E). Transcription of the fusion was up-regulated in this strain, by factors increasing from 1.3 to 2.3 over time, indicating that Spo0A negatively affects *krs* transcription. To discriminate between direct and indirect repression, we compared *lacZ* expression generated by the FL fusion and a fusion covering the same region but where the putative Spo0A box was mutated from TTTGACGAA to TTTcAgGAA to prevent Spo0A binding (fusion 0A*) [24]. Transcription from this fusion was not significantly higher than that from the FL fusion (Fig. 3D). Although we cannot rule out a possible role for the Spo0A box, our results suggest that the effect of Spo0A on *krs* transcription is indirect.

krs *expression responds to growth conditions*

The involvement of Spo0A in the expression of the *krs* operon prompted us to assess if the transcription of this locus was affected by various environmental conditions. We constructed a strain harboring the chromosomal transcriptional fusion *krsC*::*lacZ* to measure its activity in different media and conditions. We first measured *krs* transcription in shaken liquid LB, BHI, LB-peptone (LBP) or HCT. All four media are rich and complex, however LBP and HCT, unlike LB and BHI, are known to promote efficient *B. thuringiensis* differentiation like sporulation and biofilm formation [22,23]. Interestingly, we could show that *krs* transcription exhibited different patterns in LB and BHI *versus* LBP and HCT (Fig. 4A). Transcription started 1-2 hours earlier in HCT and LBP than in LB and BHI, respectively. In HCT and LBP, maximal β-galactosidase activity was measured at t3, while in LB and BHI β-galactosidase activity increased at least up to t5. The delayed expression of the *krs* operon in LB medium compared to HCT is in agreement with its activation by NprR, as previously described for *nprA*, another NprR-regulated gene [28].

To gain deeper insight into *krs* transcription, we used *gfp* as a reporter to assess *krs* transcription at the single cell level. We used a transcriptional fusion between the full-length *krs* promoter region and the s*gfp* reporter gene on the pHT304 plasmid to overcome low fluorescence issues. We measured fluorescence levels of a *B. thuringiensis* strain containing our plasmid in shaken LB and in a colony grown on LB agar (Fig. 4B). In shaken medium, the fluorescence signal was maximal at t20 (data not shown) and was evenly distributed around a single median, reflecting a uni-modal distribution of *krs* transcription (Fig. 4B, top panel). In a colony grown on LB agar, the fluorescence signal was maximal 72 hours post-inoculation (data not shown) and was distributed around two medians, reflecting a bi-modal transcription pattern (Fig. 4B, bottom panel). The leftmost peak overlaps with background auto-fluorescence, suggesting that the colony contains two distinct populations, of which only one transcribes *krs* at relatively homogeneous levels. This heterogeneity could result from the differential expression of *krs* in cells at different developmental stages within the biofilm.

The krs *operon is not essential for necrotrophism*

It was initially published that kurstakin was essential for the survival of *B. thuringiensis* in the insect cadaver [13], however we recently showed that it was incorrect and resulted from a mistake (erratum pending). The *krs* locus was supposed to be deleted from the *B. thuringiensis* 407 strain, but instead was mistakenly deleted from *B. cereus* ATCC 14579, a strain in which the *nprR* gene is interrupted by a transposon. As NprR was shown to be required for necrotrophism, the loss of necrotrophism of the *B. cereus* ATCC 14579 Δ*krsABC* strain compared to the *B. thuringiensis* 407 WT strain most likely results from the absence of NprR. Indeed, new results indicate that a *B. thuringiensis* Δ*krsC* mutant, unable to produce kurstakin (P. Jacques, personal communication), was actually not affected in its necrotrophic ability (Fig. S5). However, repeated experiments (data not shown) confirm that constitutive expression of the *krs* operon in a Δ*nprR-nprX* mutant partially restores the survival of bacteria in the insect cadaver, as it was initially shown [13]. Together, these results suggest that kurstakin could play a non-essential role in necrotrophism. Unfortunately we could not complement *krsC* deletion with our methodology, due to the important length of this gene (14.880 nt).

The krs *operon is involved in biofilm formation*

The involvement of the *krs* operon in biofilm formation was already investigated by Dubois and coll. but suffered the same strain mistake described in the previous paragraph [13]. However, the potential biophysical properties of kurstakin and its probable secretion *via* KrsE suggested that this lipopeptide could still play a role in biofilm formation [5,14,15]. We compared biofilm formation of the *B. thuringiensis* 407 *ΔkrsC* strain to that of the WT strain in various conditions. Absorbance measures showed that biofilms formed by the Δ*krsC* mutant strain in glass tubes were 40% less dense than the WT strain after 48 hours, while biofilms formed in PVC plate wells had the same density after 24 hours (Fig. 5A). However, surface pellicles formed by the Δ*krsC* mutant strain in the PVC plate wells exhibited an ultra-structure different from that of the WT strain (Fig.

5B). Pellicles formed by the WT strain seemed organized in dense networks of alveoli while pellicles formed by the Δ*krsC* mutant strain was made of larger petals loosely organized. These pellicles were also different in their resistance to pressure. We used a homemade sealed tip Pasteur pipette to apply pressure on top of the pellicles (Movie S1). While we could stab the pellicle formed by the WT strain, the Δ*krsC* pellicle sunk at the bottom of the well, as if the Δ*krsC* pellicle was more loosely attached to the surrounding material than its WT equivalent. Confocal microscopy observations of biofilms formed on filters floating at the surface of static cultures shows that the Δ*krsC* mutant biofilm architecture differed from that of the WT strain (Fig. 5C). Organization of the cells in the WT pellicle appears homogeneous and compact, while cells seem to form scattered patches of various densities in the Δ*krsC* pellicle. Most of the time, Δ*krsC* pellicles were also much thinner and contained much fewer cells than WT ones (Fig. S5).

Discussion

Kurstakin belongs to a new class of lipopeptides that has been recently discovered. Transcription of the *krs* locus, responsible for kurstakin synthesis, is activated at the onset of sporulation by NprR~NprX, a quorum-sensing system involved in the necrotrophism of *B. thuringiensis* [13]. Here we report additional data on the transcriptional regulation of this locus and describe a function of this lipopeptide in a model *B. thuringiensis* strain.

The *krs* locus contains the genes encoding the three kurstakin synthetases (*krsA*, *krsB* and *krsC*), a phosphopatheteinyl transferase (*sfp*), a type II thioesterase (*krsD*) and a membrane protein (*krsE*) potentially involved in the export of kurstakin [14,15]. Considering the function of other NRPS loci, Sfp and KrsD are thought to be involved in the turn-over of the kurstakin synthetases as phosphopantheteinyl transferase and type II thioesterase, respectively [49]. The *krs* locus was only found in genomes of the *B. cereus* group, revealing that kurstakin could be specific to this group. Interestingly, *krs* was found mostly in the *B. thuringiensis* species. This could imply that kurstakin is involved in the ecology of this organism, potentially in the infectious cycle of host larvae. Two complete *krs* loci were also found on big plasmids, raising the possibility for *krs* to be transferred between strains.

The *krs* locus consists of 6 genes spanning 30 kb. We showed that all 6 genes in the *krs* locus form an operon and we identified a transcription start site upstream of *krsE*, 327 bp upstream of its translation start codon. Analysis of *krs* expression kinetics in various media indicated that the transcription of this locus was influenced by environmental cues. We also showed that *krs* transcription was negatively affected by Spo0A, probably indirectly. However we were unable to show a regulation of *krs* expression by SinR, either by monitoring *krs* expression in a *B. thuringiensis* 407 Δ*sinIR* strain (Fig. S4) or by comparing a reference P_{krs}'-*lacZ* fusion to a fusion lacking both putative SinR boxes (Fig. 3D). These results do not corroborate microarray analyses reporting a possible repression of *krs* by SinR in LBP medium [16]. This discrepancy might result from the different conditions and methodologies used in these two studies.

We found *krs* to be transcribed earlier in media favorable to sporulation than in rich sporulation-delaying media. This is in agreement with the regulation of *krs* transcription by NprR and it is notable that the expression of *nprA*, another NprR-regulated gene, was up-regulated in HCT compared to LB [28].

We also showed that a sequence encompassing the DNA palindrome spanning the -497 to -207 region, relative to the TSS, negatively affected *krs* expression. This palindrome was not found elsewhere in the genome of *B. thuringiensis* 407, even with permissive settings. This suggests that regulation from this element is specific to *krs*, though we cannot rule out the presence in the genome of different sequences exhibiting the same function. Whether this region requires a *trans* regulatory element or acts in *cis* remains to be determined. Recent RNA-seq data from our laboratory (unpublished) confirm the position of the TSS and show no apparent transcription of the region including the palindrome, ruling out the hypotheses of a small antisense RNA or a riboswitch. Further studies are required to precisely determine the molecular function of the palindrome, beginning by the determination of a potential protein-DNA interaction.

The product of the *krs* locus, kurstakin, could be involved at different stages of the life cycle of *B. thuringiensis*. Here we showed that *krs* was required for proper structuration and resistance of *B. thuringiensis* biofilms. When formed in the absence of *krs* transcription, the biofilm exhibited a different ultra- and micro-structure compared to that of the WT strain and the pellicle behaved differently when subjected to pressure. Two non-mutually exclusive hypotheses could explain this phenotype: the Δ*krsC* pellicle could be less adherent to the surrounding surface, and/or it could be more robust than the WT strain. Both cases would produce the same outcome: the forces attaching the pellicle to the surrounding material are lower than the forces structuring the pellicle, which is precisely what we observe when applying on the pellicle. It should be noted that the involvement of the *krs* operon in biofilm formation was investigated only in *B. thuringiensis* 407. Some strains lacking the operon are known to form biofilms, and it would be interesting to compare their structure to biofilms formed by *B. thuringiensis* 407 WT and Δ*krsC*.

Extracting and purifying kurstakin would allow further experiments. For instance, rheology studies on biofilms with purified kurstakin could help assess its precise implication in the biofilm formation process. Exogenous purified kurstakin could also be used to try to rescue the phenotype of a biofilm formed by the kurstakin-defective mutant.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

We are indebted to Agnès Réjasse, Francis Repoila and Françoise Wessner for their help setting up the RNA purification protocol. We are grateful to Stéphane Perchat and Rafael Patiño-Navarrete for fruitful discussions. We thank the referees for thoughtful advice. This study was funded by the French “Agence Nationale de la Recherche” (Bt-Surf, ANR-12-EMMA-0005). S. Gélis-Jeanvoine was supported by a fellowship from the French “Ministère de l’agriculture”.

References

[1] Jacques P. Surfactin and other lipopeptides from *Bacillus* spp. In: Soberón-Chávez G, editor. Biosurfactants, vol. 20, Berlin, Heidelberg: Springer Berlin Heidelberg; 2011, p. 57–91. doi:10.1007/978-3-642-14490-5.

[2] Cameotra SS, Makkar RS. Recent applications of biosurfactants as biological and immunological molecules. Curr Opin Microbiol 2004;7:262–6. doi:10.1016/j.mib.2004.04.006.

[3] Kanlayavattanakul M, Lourith N. Lipopeptides in cosmetics. Int J Cosmet Sci 2010;32:1–8. doi:10.1111/j.1468-2494.2009.00543.x.

[4] Khong NG, Randoux B, Tayeh C, Coutte F, Bourdon N, Tisserant B, *et al*. Induction of resistance in wheat against powdery mildew by bacterial cyclic lipopeptides. Commun Agric Appl Biol Sci 2012;77:39–51.

[5] Raaijmakers JM, De Bruijn I, Nybroe O, Ongena M. Natural functions of lipopeptides from *Bacillus* and *Pseudomonas*: more than surfactants and antibiotics. FEMS Microbiol Rev 2010;34:1037–62. doi:10.1111/j.1574-6976.2010.00221.x.

[6] Magnuson R, Solomon J, Grossman AD. Biochemical and genetic characterization of a competence pheromone from *B. subtilis*. Cell 1994;77:207–16.

[7] Nakano MM, Magnuson R, Myers A, Curry J, Grossman AD, Zuber P. *srfA* is an operon required for surfactin production, competence development, and efficient sporulation in *Bacillus subtilis*. J Bacteriol 1991;173:1770–8.

[8] López D, Vlamakis H, Losick R, Kolter R. Paracrine signaling in a bacterium. Genes Dev 2009;23:1631–8. doi:10.1101/gad.1813709.

[9] López D, Kolter R. Extracellular signals that define distinct and coexisting cell fates in *Bacillus subtilis*. FEMS Microbiol Rev 2010;34:134–49. doi:10.1111/j.1574-6976.2009.00199.x.

[10] Vilas-Bôas GT, Peruca APS, Arantes OMN. Biology and taxonomy of *Bacillus cereus*, *Bacillus anthracis*, and *Bacillus thuringiensis*. Can J Microbiol 2007;53:673–87. doi:10.1139/W07-029.

[11] Hathout Y, Ho YP, Ryzhov V, Demirev P, Fenselau C. Kurstakins: a new class of lipopeptides isolated from *Bacillus thuringiensis*. J Nat Prod 2000;63:1492–6.

[12] Bumpus SB, Evans BS, Thomas PM, Ntai I, Kelleher NL. A proteomics approach to discovering natural products and their biosynthetic pathways. Nat Biotechnol 2009;27:951–6. doi:10.1038/nbt.1565.

[13] Dubois T, Faegri K, Perchat S, Lemy C, Buisson C, Nielsen-LeRoux C, *et al*. Necrotrophism is a quorum-sensing-regulated lifestyle in *Bacillus thuringiensis*. PLoS Pathog 2012;8:e1002629. doi:10.1371/journal.ppat.1002629.

[14] Béchet M, Caradec T, Hussein W, Abderrahmani A, Chollet M, Leclère V, *et al*. Structure, biosynthesis, and properties of kurstakins, nonribosomal lipopeptides from *Bacillus* spp. Appl Microbiol Biotechnol 2012;95:593–600. doi:10.1007/s00253-012-4181-2.

[15] Li X, Yang H, Zhang D, Li X, Yu H, Shen Z. Overexpression of specific proton motive force-dependent transporters facilitate the export of surfactin in *Bacillus subtilis*. J Ind Microbiol Biotechnol 2015;42:93–103. doi:10.1007/s10295-014-1527-z.

[16] Fagerlund A, Dubois T, Økstad O-A, Verplaetse E, Gilois N, Bennaceur I, *et al*. SinR Controls Enterotoxin Expression in *Bacillus thuringiensis* Biofilms. PLoS One 2014;9:e87532. doi:10.1371/journal.pone.0087532.

[17] Mazzola M, de Bruijn I, Cohen MF, Raaijmakers JM. Protozoan-induced regulation of cyclic lipopeptide biosynthesis is an effective predation defense mechanism for *Pseudomonas fluorescens*. Appl Environ Microbiol 2009;75:6804–11. doi:10.1128/AEM.01272-09.

[18] Lereclus D, Arantes O, Chaufaux J, Lecadet M. Transformation and expression of a cloned delta-endotoxin gene in *Bacillus thuringiensis*. FEMS Microbiol Lett 1989;51:211–7.

[19] Bertani G. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. J Bacteriol 1951;62:293–300.

[20] Rosenow EC. Studies on Elective Localization Focal Infection with Special Reference to Oral Sepsis. J Dent Res 1919;1:205–67. doi:10.1177/00220345190010030101.

[21] Haden RL. Elective localization in the eye of bacteria from infected teeth. Arch Intern Med 1923;32:828. doi:10.1001/archinte.1923.00110240019002.

[22] Lecadet MM, Blondel MO, Ribier J. Generalized transduction in *Bacillus thuringiensis* var. berliner 1715 using bacteriophage CP-54Ber. J Gen Microbiol 1980;121:203–12.

[23] Auger S, Krin E, Aymerich S, Gohar M. Autoinducer 2 affects biofilm formation by *Bacillus cereus*. Appl Environ Microbiol 2006;72:937–41. doi:10.1128/AEM.72.1.937-941.2006.

[24] York K, Kenney TJ, Satola S, Moran CP, Poth H, Youngman P. Spo0A controls the sigma A-dependent activation of *Bacillus subtilis* sporulation-specific transcription unit *spoIIE*. J Bacteriol 1992;174:2648–58.

[25] Lereclus D, Vallade M, Chaufaux J, Arantes O, Rambaud S. Expansion of insecticidal host range of *Bacillus thuringiensis* by *in vivo* genetic recombination. Bio/Technology 1992;10:418–21.

[26] Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, *et al*. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol 2006;7:3. doi:10.1186/1471-2199-7-3.

[27] Cadot C, Tran S-L, Vignaud M-L, De Buyser M-L, Kolstø A-B, Brisabois A, *et al*. InhA1, NprA, and HlyII as candidates for markers to differentiate pathogenic from nonpathogenic *Bacillus cereus* strains. J Clin Microbiol 2010;48:1358–65. doi:10.1128/JCM.02123-09.

[28] Perchat S, Dubois T, Zouhir S, Gominet M, Poncet S, Lemy C, *et al*. A cell-cell communication system regulates protease production during sporulation in bacteria of the

Bacillus cereus group. Mol Microbiol 2011;82:619–33. doi:10.1111/j.1365-2958.2011.07839.x.

[29] Houry A, Briandet R, Aymerich S, Gohar M. Involvement of motility and flagella in *Bacillus cereus* biofilm formation. Microbiology 2010;156:1009–18. doi:10.1099/mic.0.034827-0.

[30] Vlamakis H, Aguilar C, Losick R, Kolter R. Control of cell fate by the formation of an architecturally complex bacterial community. Genes Dev 2008;22:945–53. doi:10.1101/gad.1645008.

[31] Verplaetse E, Slamti L, Gohar M, Lereclus D. Cell Differentiation in a *Bacillus thuringiensis* Population during Planktonic Growth, Biofilm Formation, and Host Infection. MBio 2015;6:e00138–15. doi:10.1128/mBio.00138-15.

[32] Sheppard AE, Poehlein A, Rosenstiel P, Liesegang H, Schulenburg H. Complete Genome Sequence of *Bacillus thuringiensis* Strain 407 Cry-. Genome Announc 2013;1. doi:10.1128/genomeA.00158-12.

[33] Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA, *et al*. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. Nucleic Acids Res 2011;39:W339–46. doi:10.1093/nar/gkr466.

[34] Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, *et al*. antiSMASH 3.0--a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res 2015:gkv437 – . doi:10.1093/nar/gkv437.

[35] Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, *et al*. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25:3389–402.

[36] Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, *et al*. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 2011;7:539. doi:10.1038/msb.2011.75.

[37] Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol 2000;17:540–52.

[38] Talavera G, Castresana J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol 2007;56:564–77. doi:10.1080/10635150701472164.

[39] Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol 2009;26:1641–50. doi:10.1093/molbev/msp077.

[40] Price MN, Dehal PS, Arkin AP. FastTree 2--approximately maximum-likelihood trees for large alignments. PLoS One 2010;5:e9490. doi:10.1371/journal.pone.0009490.

582 [41] Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2--a
583 multiple sequence alignment editor and analysis workbench. Bioinformatics 2009;25:1189–
584 91. doi:10.1093/bioinformatics/btp033.

585 [42] Van Ert MN, Easterday WR, Huynh LY, Okinaka RT, Hugh-Jones ME, Ravel J, *et al*. Global
586 Genetic Population Structure of *Bacillus anthracis*. PLoS One 2007;2:e461.
587 doi:10.1371/journal.pone.0000461.

588 [43] Huang L, Cheng T, Xu P, Cheng D, Fang T, Xia Q. A genome-wide survey for host response
589 of silkworm, *Bombyx mori* during pathogen *Bacillus bombyseptieus* infection. PLoS One
590 2009;4:e8098. doi:10.1371/journal.pone.0008098.

591 [44] Guinebretière M-H, Thompson FL, Sorokin A, Normand P, Dawyndt P, Ehling-Schulz M, *et*
592 *al*. Ecological diversification in the *Bacillus cereus* Group. Environ Microbiol 2008;10:851–
593 65. doi:10.1111/j.1462-2920.2007.01495.x.

594 [45] Tourasse NJ, Okstad OA, Kolstø A-B. HyperCAT: an extension of the SuperCAT database
595 for global multi-scheme and multi-datatype phylogenetic analysis of the *Bacillus cereus*
596 group population. Database (Oxford) 2010;2010:baq017. doi:10.1093/database/baq017.

597 [46] Moran CP, Lang N, LeGrice SF, Lee G, Stephens M, Sonenshein AL, *et al*. Nucleotide
598 sequences that signal the initiation of transcription and translation in *Bacillus subtilis*. Mol
599 Gen Genet 1982;186:339–46.

600 [47] Chu F, Kearns DB, Branda SS, Kolter R, Losick R. Targets of the master regulator of biofilm
601 formation in *Bacillus subtilis*. Mol Microbiol 2006;59:1216–28. doi:10.1111/j.1365-
602 2958.2005.05019.x.

603 [48] Spiegelman GB, Bird TH, Voon Y. Transcription regulation by the *Bacillus subtilis* response
604 regulator Spo0A. In: Hoch JA, Silhavy TJ, editors. Two Compon. signal Transduct. American
605 S, Washington D.C.: 1995, p. 159–79.

606 [49] Finking R, Marahiel MA. Biosynthesis of nonribosomal peptides. Annu Rev Microbiol
607 2004;58:453–88. doi:10.1146/annurev.micro.58.030603.123615.

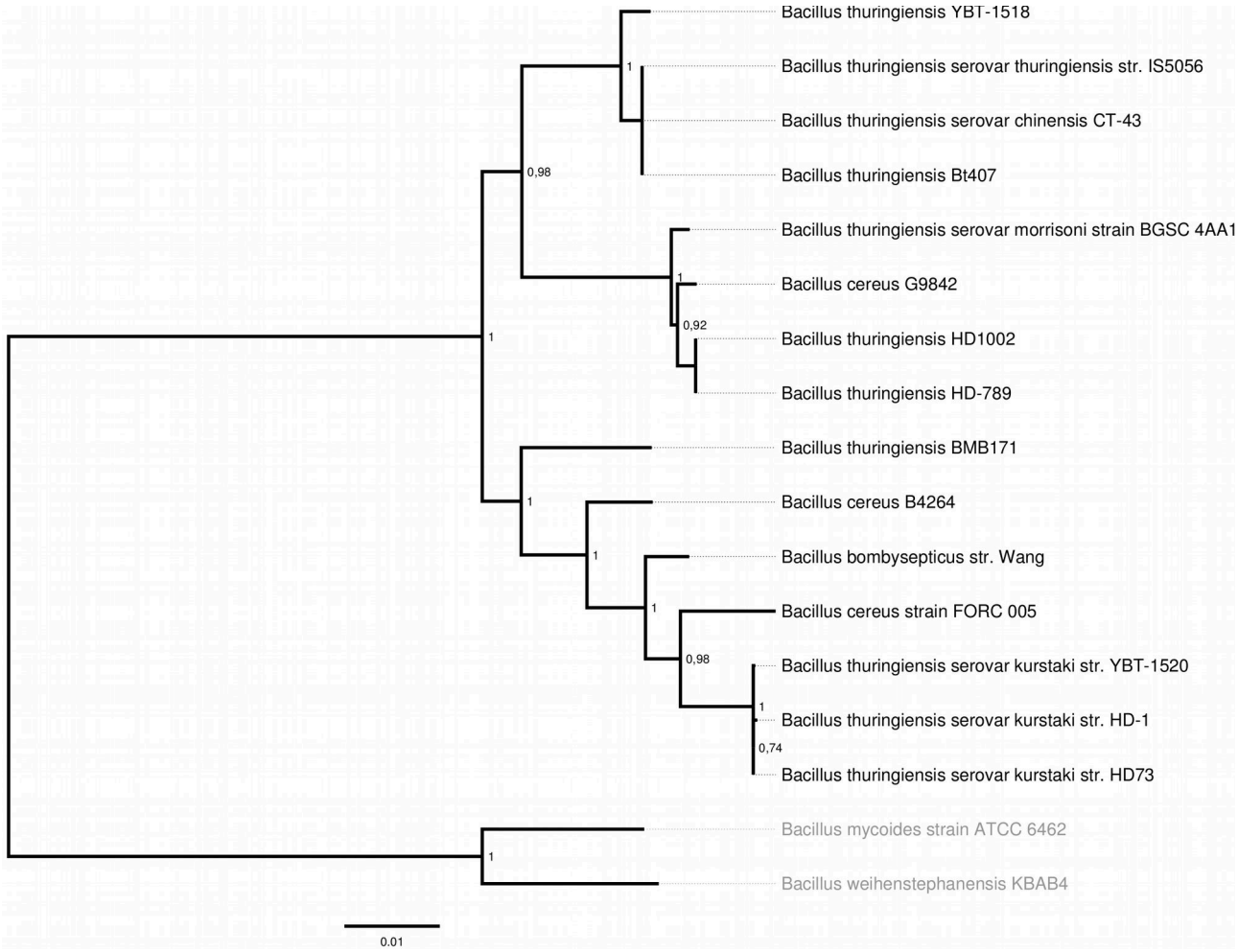
608 [50] Lereclus D, Agaisse H, Gominet M, Chaufaux J. Overproduction of Encapsulated
609 Insecticidal Crystal Proteins in a *Bacillus thuringiensis spo0A* Mutant. Bio/Technology
610 1995;13:67–71. doi:10.1038/nbt0195-67.

611 [51] Agaisse H, Lereclus D. Structural and functional analysis of the promoter region involved in
612 full expression of the *cryIIIA* toxin gene of *Bacillus thuringiensis*. Mol Microbiol
613 1994;13:97–107.

614 [52] Lauderdale KJ, Malone CL, Boles BR, Morcuende J, Horswill AR. Biofilm dispersal of
615 community-associated methicillin-resistant *Staphylococcus aureus* on orthopedic implant
616 material. J Orthop Res 2010;28:55–61. doi:10.1002/jor.20943.

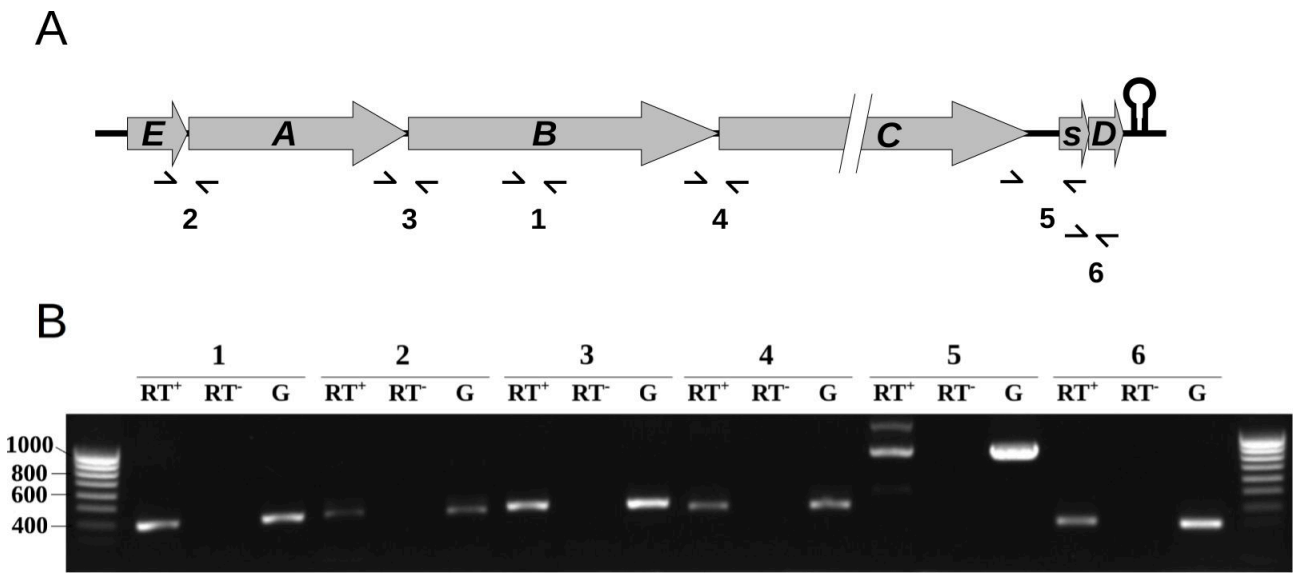
Figures

Figure 1. Krs^{+} strains mainly belong to cluster II.



The sequences of the products of the six *krs* genes were concatenated for each strain and aligned with ClustalΩ. This global alignment was used by FastTree to build the approximately-maximum-likelihood tree of all Krs^{+} loci. Branch supports are indicated. Strains in black belong to clade II, strains in gray belong to clade III.

Figure 2. RT-PCR analysis of the *krs* locus.

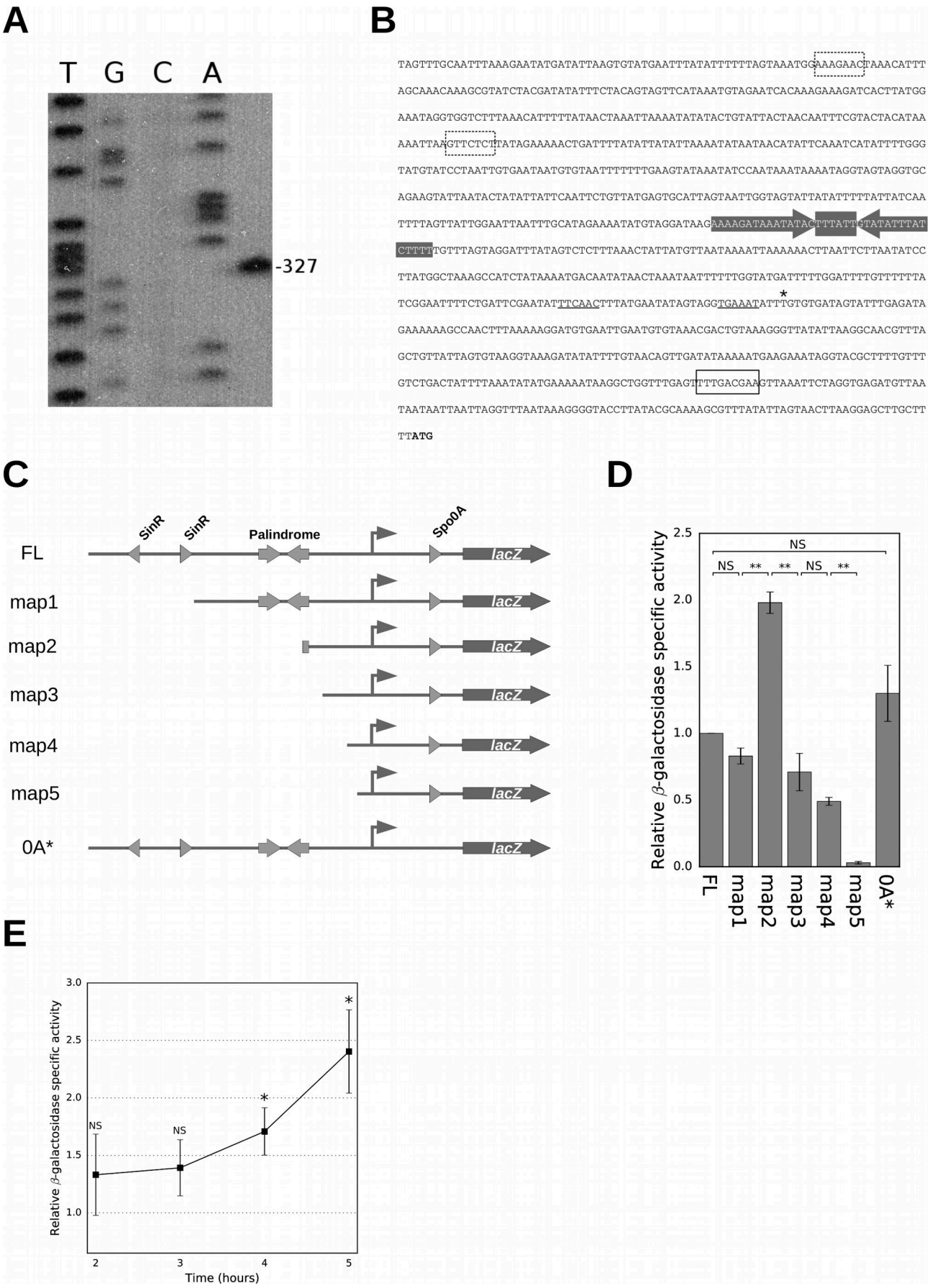


A. Schematic representation to scale of the *krs* locus of *B. thuringiensis* 407. *E*: *krsE*, *A*: *krsA*, *B*: *krsB*, *C*: *krsC*, *s*: *sfp*, *D*: *krsD*. Note the scale break in *krsC*. The stem-loop structure in 3' represents a putative transcription terminator. Below the locus are represented the primers pairs used in the RT-PCR (see experimental procedures). **B.** Photograph after electrophoretic migration of the RT-PCR products in a 0.8% agarose gel containing ethidium bromide. Each primer pair was used in a reaction with or without reverse transcriptase (RT^+ or RT^-, respectively) using cDNA synthesized from total RNA extracted from *B. thuringiensis* 407 as a template, and a reaction using *B. thuringiensis* 407 genomic DNA as a template (G).

638 Figure 3. Mapping of the *krs* promoter region.

639

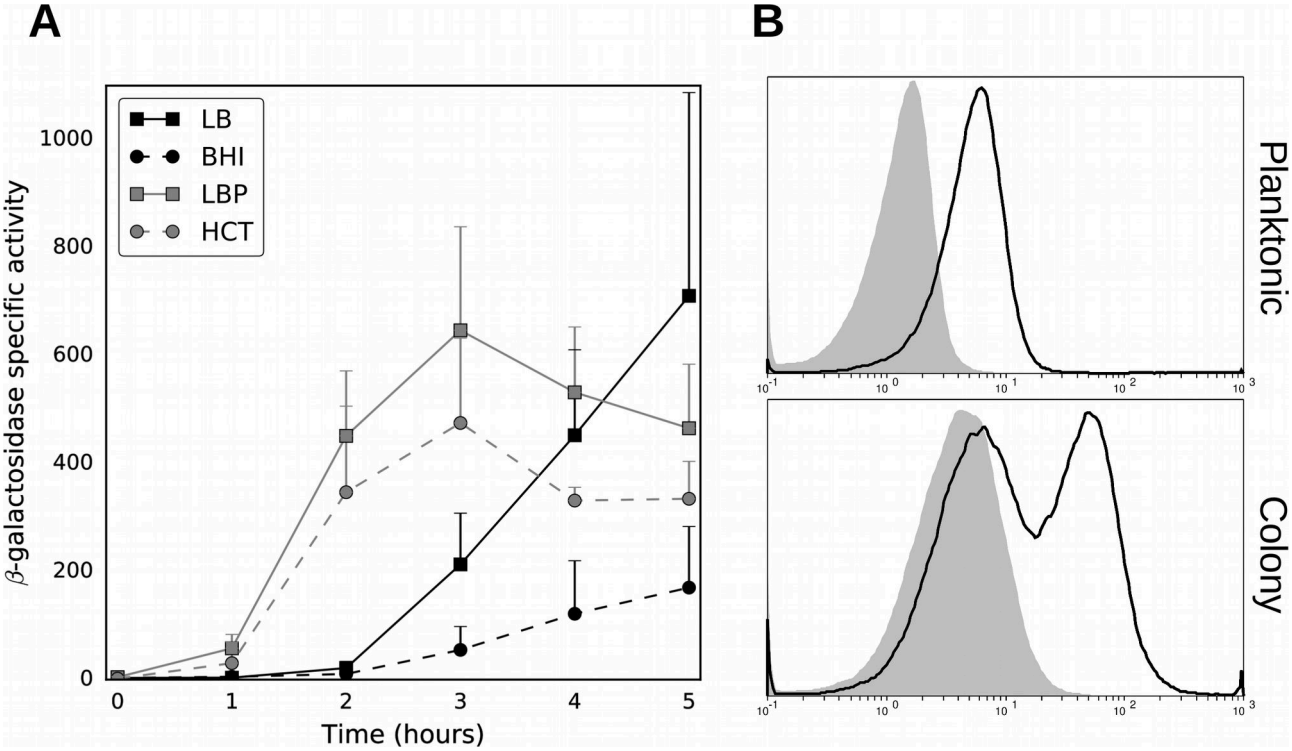
640



A. Total RNA was subjected to primer extension analysis using oligonucleotide tss_bc2450 complementary to the 5' end of the *krsE* gene. The same oligonucleotide was used to prime dideoxy sequencing reactions from the *B. thuringiensis* 407 chromosome (lanes T, G, C, A). Position of the transcription start site is given with reference to the translation start codon of *krsE*. **B.** Sequence of the intergenic space upstream of *krsE*. The two putative SinR binding sites are framed (dashes). The putative Spo0A binding site is framed (plain). The palindrome is highlighted in gray. The TSS is marked by a star. The putative -35 and -10 boxes are underlined and the putative ATG of *krsE* is in bold characters. **C.** Schematic representation of the different fragments that were used in transcriptional fusions with the *lacZ* reporter gene on the pHT304-18Z plasmid **D.** Relative β-galactosidase specific activity generated by the different constructions described in B, at t3 in HCT at 30°C, normalized to the activity of the FL reference construct. NS: not significant, **: $p < 0.005$. **E.** Relative β-galactosidase specific activity generated by the FL construct in the Δ*spo0A* strain in HCT at 30°C, normalized to its activity in the WT strain. Measures were taken hourly starting at t2. NS: not significant, *: $p < 0.05$.

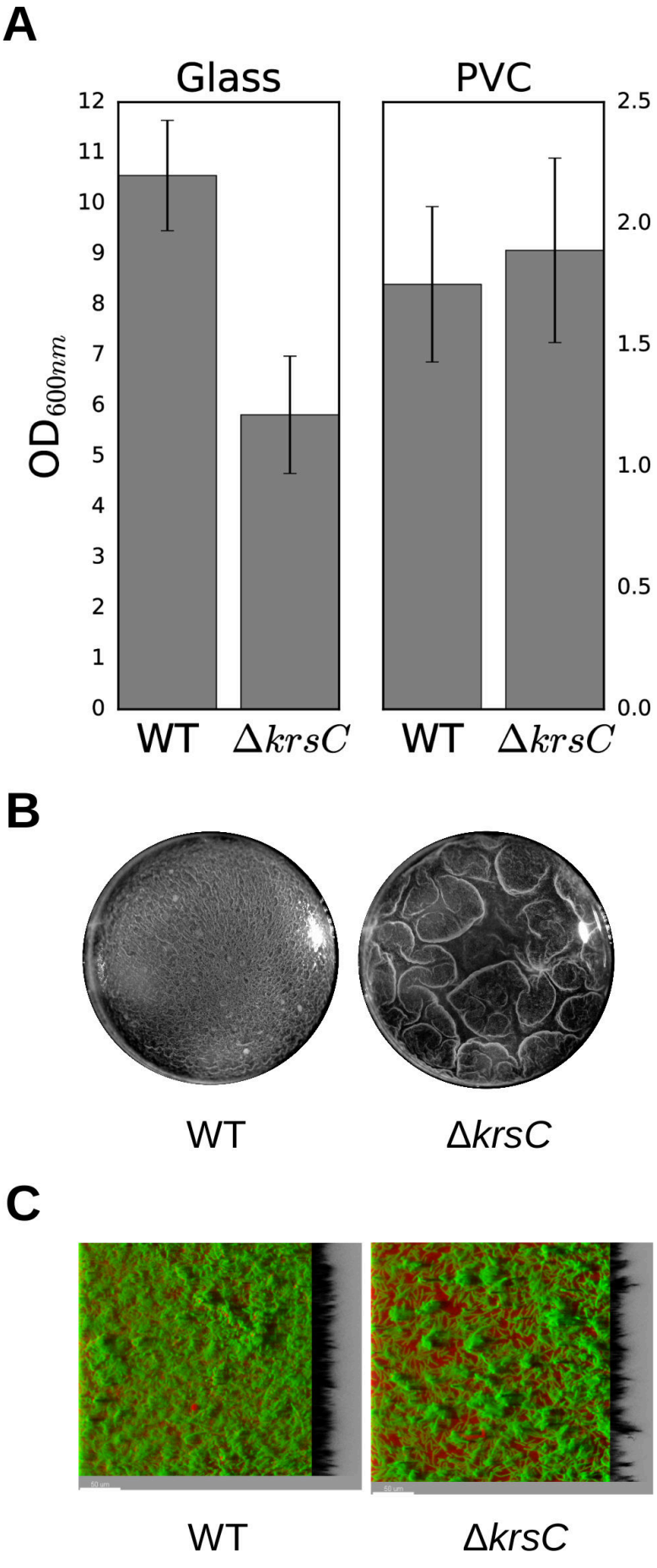
655 Figure 4. *krs* transcription is sensitive to growth conditions.

656



658 **A.** β-galactosidase specific activity of the *krsC*::*lacZ* chromosomal fusion in shaken LB, BHI, LBP
659 and HCT at 30°C. Time 0 marks the beginning of the transition between exponential and stationary
660 phase. Error bars were plotted asymmetrically for the sake of readability. **B.** Flow cytometry
661 analysis of the fluorescence distribution in cells harboring the pHT304ΩP$_{krs}$(FL)'-*sgfp* plasmid in
662 shaken LB (upper panel) and a colony grown on LB agar (lower panel). Samples were taken 20
663 hours after the end of stationary phase in shaken LB, and 72 hours after colony inoculation on LB
664 agar. Auto-fluorescence of a reporter-less strain is represented by the gray area, fluorescence of the
665 bacteria harboring the plasmid is represented in black outline. This experiment is representative of
666 at least 4 independent replicates.

667 Figure 5. Kurstakin is required for proper biofilm structuration.



A. Biofilm density in glass tubes after a 48 hours incubation and PVC plates after a 24 hour incubation in HCT at 30°C. **B.** Air-liquid pellicles of biofilms formed by the WT and Δ*krsC* strains in PVC-plate wells, observed from above with a stereo microscope **C.** 3D reconstructions of confocal laser scanning spectrum micrographs.

672 **Tables**

673 Table 1. *krs* is not evenly distributed within the *Bacillus cereus* group

674

Species of the *B. cereus* group	Frequency of Krs^+ genomes
Bacillus anthracis	0/31
Bacillus cereus	3/26
Bacillus cytotoxicus	0/1
Bacillus mycoides	1/2
Bacillus thuringiensis	12/23 (10/21)*
Bacillus toyonensis	0/1
Bacillus weihenstephanensis	1/2

* The frequency was corrected from 12/23 to 10/21 to take into account redundancy of three *B. thuringiensis* genomes (strains Bt407, CT-43 and IS5056).

Supplementary material for on-line submission

Table S1. *B. thuringiensis* 407 derived strains used in this study.

Strain	**Reference**
Bacillus thuringiensis str. 407 Cry$^-$ (*B. thuringiensis* 407)	Laboratory collection
B. thuringiensis 407 *krsC*::*lacZ*	This study
B. thuringiensis 407 *spo0A*::*aphA3*	[50]
B. thuringiensis 407 *sinI-sinR*::*tet*	[16]
B. thuringiensis 407 (pHT304ΩP$_{krs}$(FL)'-*sgfp*)	This study
B. thuringiensis 407 (pHT304-18ΩP$_{krs}$(FL)'-*lacZ*)	This study
B. thuringiensis 407 (pHT304-18ΩP$_{krs}$(map1)'-*lacZ*)	This study
B. thuringiensis 407 (pHT304-18ΩP$_{krs}$(map2)'-*lacZ*)	This study
B. thuringiensis 407 (pHT304-18ΩP$_{krs}$(map3)'-*lacZ*)	This study
B. thuringiensis 407 (pHT304-18ΩP$_{krs}$(map4)'-*lacZ*)	This study
B. thuringiensis 407 (pHT304-18ΩP$_{krs}$(map5)'-*lacZ*)	This study
B. thuringiensis 407 (pHT304-18ΩP$_{krs}$(ΔPal)'-*lacZ*)	This study
B. thuringiensis 407 (pHT304-18ΩP$_{krs}$(0A*)'-*lacZ*)	This study

680 Table S2. Plasmids used in this study.

681

682

Name	Features
pRN5101	Thermo-sensitive replication origin. Used in *B. thuringiensis* 407 to mediate homologous recombination with the chromosome [25].
pHT304-18Z	Four-copies plasmid containing the *lacZ* gene downstream of a multiple-cloning site [51]. Used as a reporter plasmid.
pRN5101Ω*krsC*::*lacZ*	Regions of 1 kb were amplified from *krsB* (upstream) with SG1 and SG2, and the end of *krsC* (downstream) with SG3 and SG4, on *B. thuringiensis* 407 genomic DNA, to allow homologous recombination with the chromosome. The promoterless *lacZ* gene was purified as a XbaI-KpnI fragment from the pHT304-18Z plasmid and the three fragments were cloned in a pRN5101 plasmid digested with BamHI and HindIII.
pHT304sGFP	The *sgfp* allele was amplified from pCM11 [52] with SG30 and SG31, digested with SphI and HindIII and cloned in a SphI-HindIII digested pHT304 plasmid.
pHT304ΩP_{krs}(FL)'-*sgfp*	The whole intergenic sequence upstream of *krsE*, including the first 84 nt of *krsE* (the -739 to +84 region), was amplified with SG32 and SG33 on *B. thuringiensis* 407 genomic DNA. It was then cloned in the pHT304sGFP plasmid between the PstI and SphI restriction sites. This construction was used to monitor *krs* transcription at the single-cell level.
pHT304-18ΩP_{krs}(FL)'-*lacZ*	The whole intergenic sequence upstream of *krsE*, including the first 84 nt of *krsE* (the -739 to +84 region), was amplified with SG34 and SG35 on *B. thuringiensis* 407 genomic DNA. It was then cloned in the pHT304-18Z plasmid between the BamHI and XbaI restriction sites. This construction was used to map P_{krs}.
pHT304-18ΩP_{krs}(map1)'-*lacZ*	The -497 to +84 region, lacking both putative SinR boxes, was amplified with SG66 and SG35 on *B. thuringiensis* 407 genomic DNA. It was then cloned in the pHT304-18Z plasmid between the BamHI and XbaI restriction sites. This construction was used to map P_{krs}.
pHT304-18ΩP_{krs}(map2)'-*lacZ*	This construction was described in [13] as BC2450'Z and covers the -207 to +84 region of the promoter. This construction was used to map P_{krs}.
pHT304-18ΩP_{krs}(map3)'-*lacZ*	The -163 to +84 region was amplified with

	SG84 and SG35 on *B. thuringiensis* 407 genomic DNA. It was then cloned in the pHT304-18Z plasmid between the BamHI and XbaI restriction sites. This construction was used to map P_{krs}.
pHT304-18ΩP_{krs}(map4)'-*lacZ*	The -77 to +84 region was amplified with SG83 and SG35 on *B. thuringiensis* 407 genomic DNA. It was then cloned in the pHT304-18Z plasmid between the BamHI and XbaI restriction sites. This construction was used to map P_{krs}.
pHT304-18ΩP_{krs}(map5)'-*lacZ*	The -45 to +84 region was amplified with SG67 and SG35 on *B. thuringiensis* 407 genomic DNA. It was then cloned in the pHT304-18Z plasmid between the BamHI and XbaI restriction sites. This construction was used to map P_{krs}.
pHT304-18ΩP_{krs}(0A*)'-*lacZ*	The pHT304-18Ω$_{Pkrs}$(FL)'-*lacZ* plasmid was used as a template in a PCR with primers SG74 and SG75, mutating the Spo0A box from TTTGACGAA to TTTCAGGAA. The whole plasmid was synthesized and matrix DNA was digested with DpnI. This construction was used to assess the effect of the putative Spo0A box on *krs* transcription.

684 Table S3. Primers used in this study.

685

686

Name	Sequence*	Restriction site
SG1	gggggatccTATGGAATTAGAAAAGGGTG	BamHI
SG2	ggtctagaCAATCAAATTACTCTCGAAT	XbaI
SG3	ggggtaccGAAATTGAATATCAGTTGAT	KpnI
SG4	gggaagcttCTACTTTTCCGTTAATGAAA	HindIII
SG30	cccgcatgcCTTAGGAGGATGATTATTTA	SphI
SG31	gggaagcttTTAGTGGTGGTGGTGGTGGTGGGATCCTTT	HindIII
SG32	ccctgcagTAGTTTGCAATTTAAAGAAT	PstI
SG33	ggggcatgcACCAAGCATAAGCAATGAAT	SphI
SG34	ctctagaTAGTTTGCAATTTAAAGAAT	XbaI
SG35	gggggatccACCAAGCATAAGCAATGAAT	BamHI
SG66	ctctagaTATAGAAAAACTGAT	XbaI
SG67	ctctagaCTGATTCGAATATTT	XbaI
SG74	CTAGAATTTAACTTCcTgAAAACTCAAACCAGC	
SG75	GCTGGTTTGAGTTTTcAgGAAGTTAAATTCTAG	
SG83	ctctagaGATTTTTGGATTTTG	XbaI
SG84	ctctagaAATAAAATTAAAAAAC	XbaI
RT1_FW	ATTGCCTCGGTTTGGAGT	
RT1_REV	ACAATTGATTCCATCCGACCT	
RT2_FW	TGGTTTTGGGGGGATTGG	
RT2_REV	GCTGCGGATGGCTTAATGG	
RT3_FW	GCAAGAGGCAGGGTTTACA	
RT3_REV	CCCCCCGAGCAAGAAACCTA	
RT4_FW	TCGCTACGCAAGCTGTTTC	
RT4_REV	ACAACGGGAGTGCTTCT	
RT5_FW	AGTGCTGAGGAAGGAAAAGAG	
RT5_REV	AGTACAAACCACCCATTCTCC	
RT6_FW	AAAAACGATTGGAAAAGGGCT	
RT6_REV	AGTTAAACGCAGATCCCCCA	
tss_bc2450	CCCTTTACAGTCGTTTACAC	

* Nucleotides not matching the template are lowercase. Restriction sites are underlined.

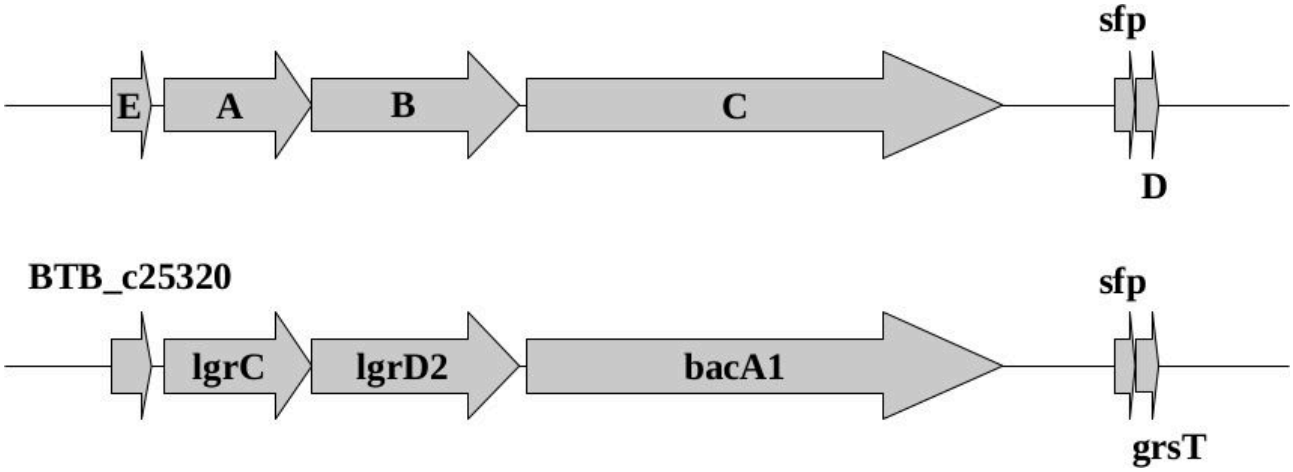
687 Table S4. Krs^{+} genomes.

688

Strain	Genbank accession
Bacillus bombysepticus str. Wang	CP007512.1
Bacillus cereus B4264	CP001176.1
Bacillus cereus G9842	CP001186.1
Bacillus cereus strain FORC_005	CP009686.1
Bacillus mycoides strain ATCC 6462	CP009691.1
Bacillus thuringiensis BMB171	CP001903.1
Bacillus thuringiensis Bt407	CP003889.1
Bacillus thuringiensis HD-789	CP003763.1
Bacillus thuringiensis HD1002	CP009351.1
Bacillus thuringiensis YBT-1518	CP005935.1
Bacillus thuringiensis serovar *chinensis* CT-43	CP001907.1
Bacillus thuringiensis serovar *kurstaki* str. HD-1	CP004870.1
Bacillus thuringiensis serovar *kurstaki* str. HD73	CP004069.1
Bacillus thuringiensis serovar *kurstaki* str. YBT-1520	CP007607.1
Bacillus thuringiensis serovar *morrisoni* strain BGSC 4AA1	CP010577.1
Bacillus thuringiensis serovar *thuringiensis* str. IS5056	CP004123.1
Bacillus weihenstephanensis KBAB4	CP000904.1

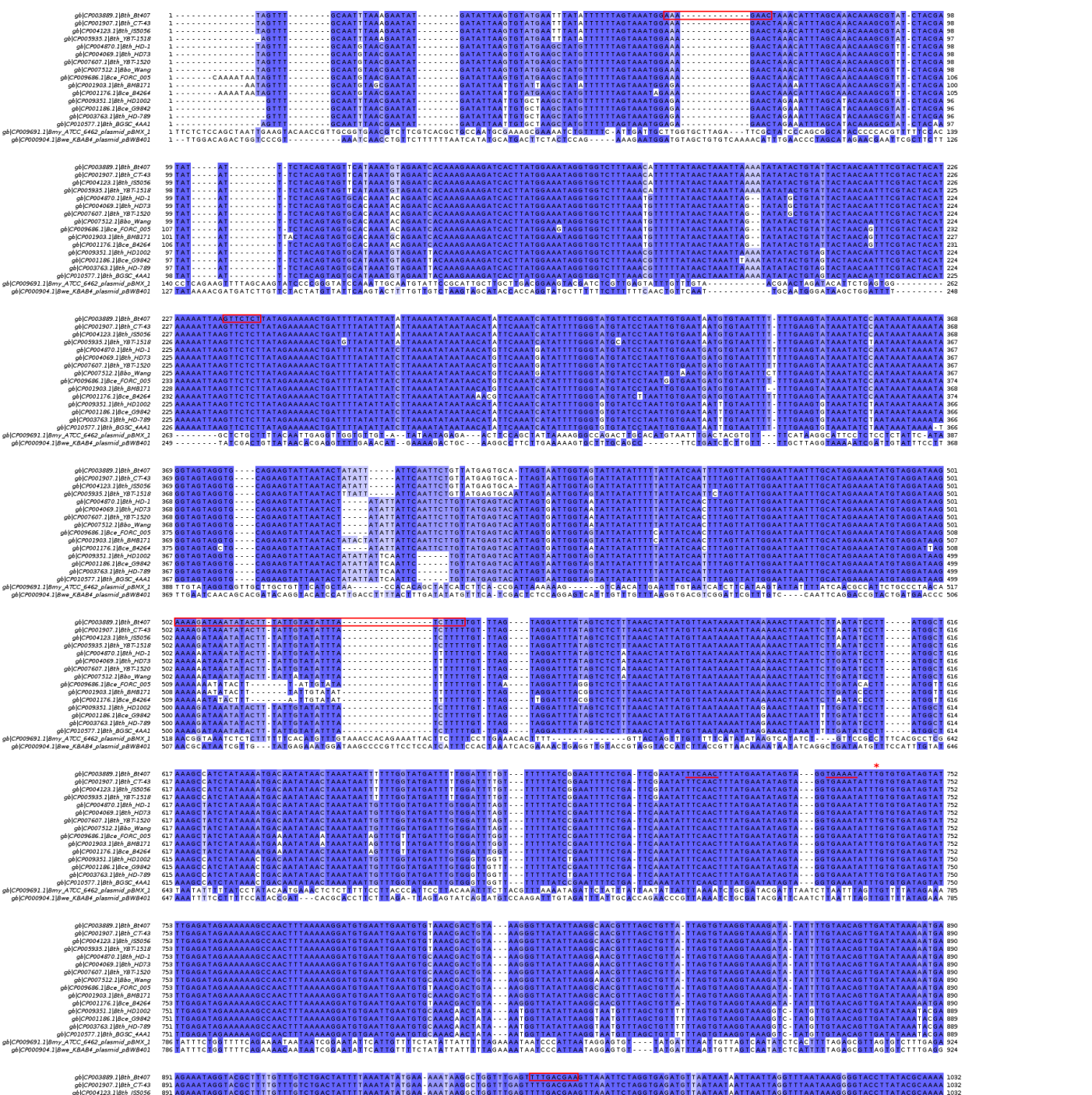
689 Figure S1. The *krs* locus in the *B. thuringiensis* 407 reference genome.

690



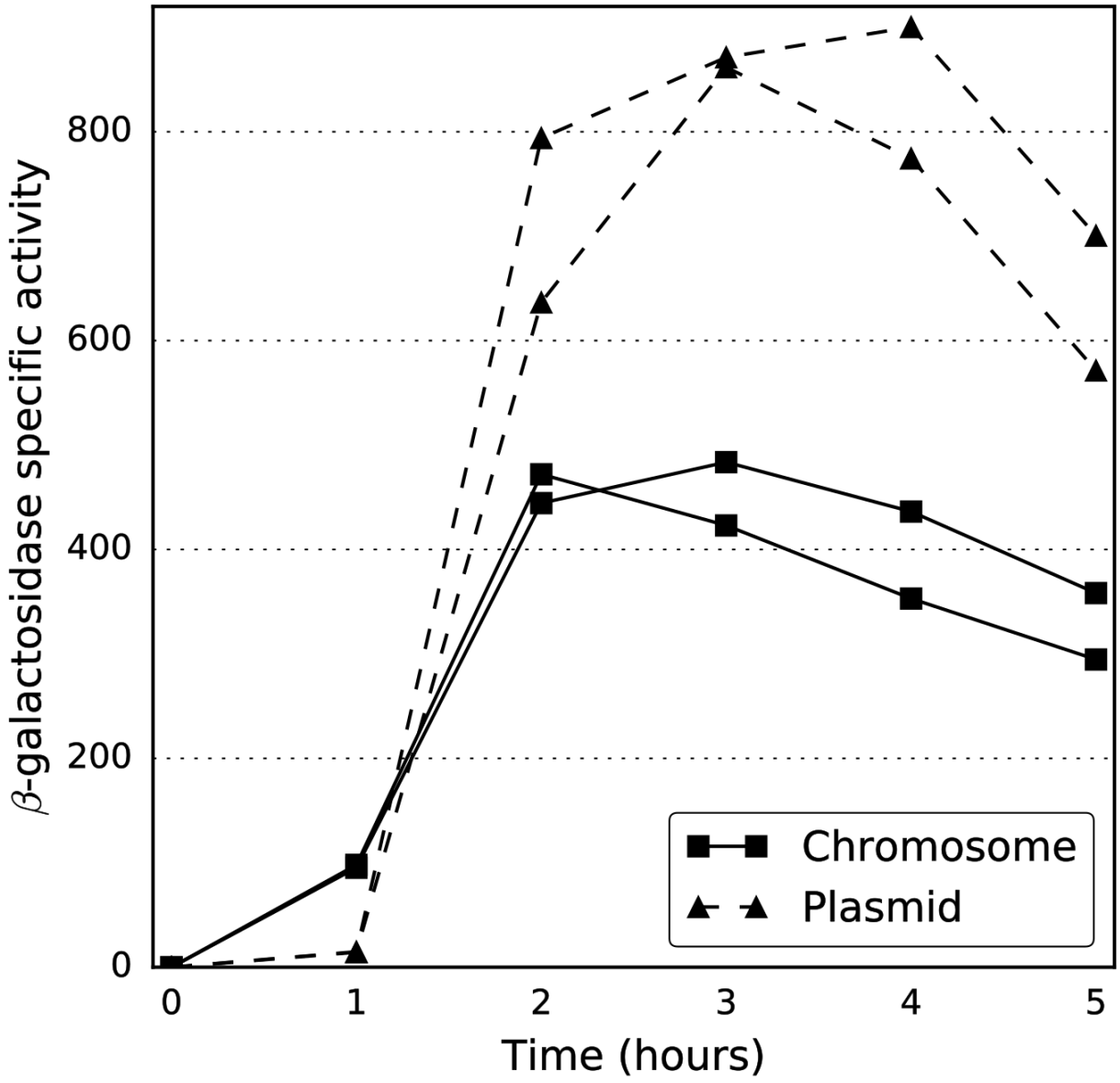
692 The top schematic represents our annotation of the *krs* locus, the bottom one represents the
693 annotation on the *B. thuringiensis* 407 reference genome [32].

Figure S2. Alignment of 17 *krs* upstream sequences.



Conservation at each position is indicated by shades of blue: the darker the more conserved.

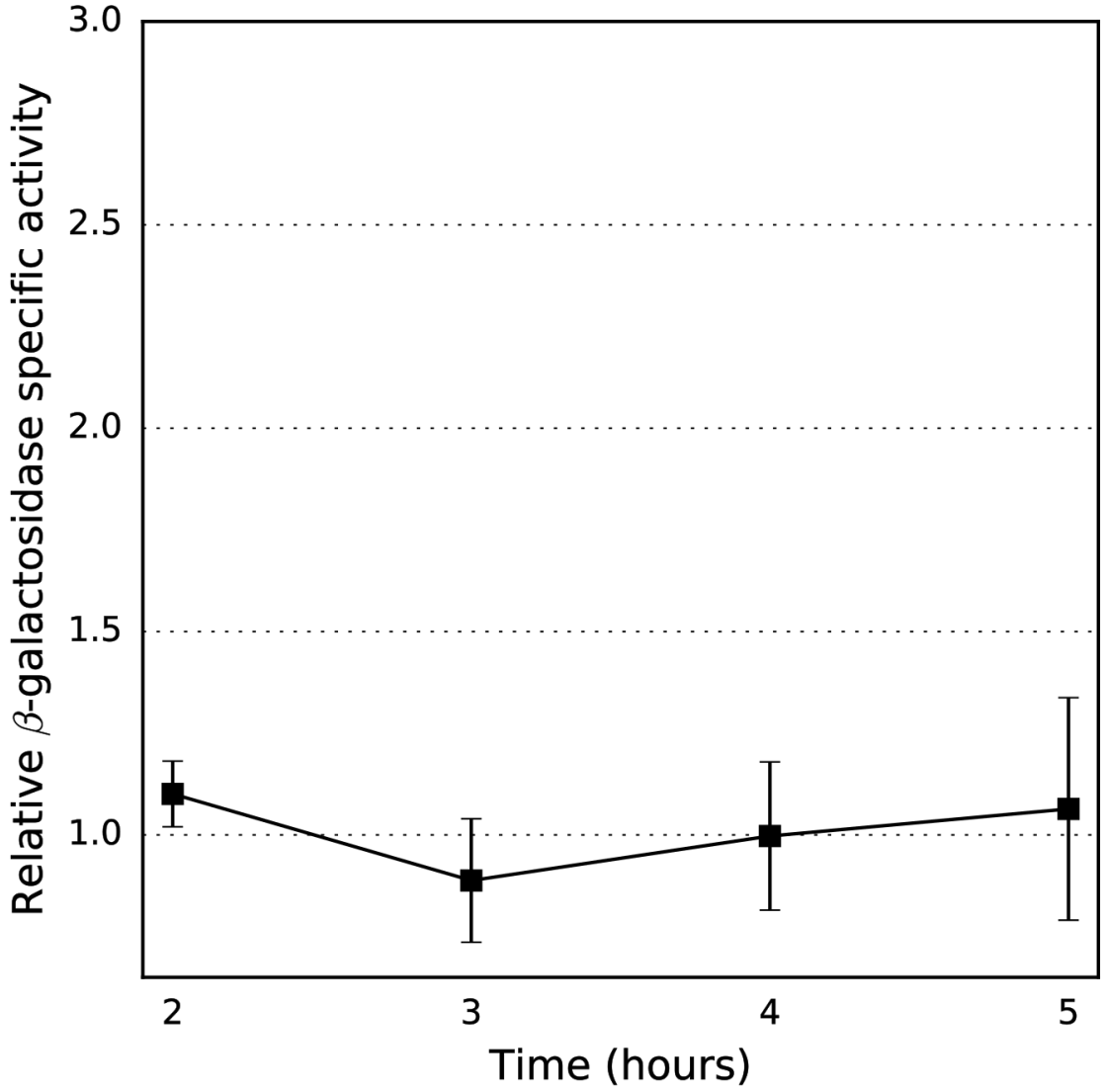
Figure S3. Plasmidic P_{krs}'-*lacZ* fusion has a higher activity than the chromosomal *krsC*::*lacZ* fusion but has a comparable pattern.



β-galactosidase specific activity generated by strains harboring chromosomal (*B. thuringiensis* 407 *krsC*::*lacZ* strain) or plasmidic (*B. thuringiensis* 407 (pHT304-18ΩP_{krs}(FL)'-*lacZ*) strain) P_{krs}'-*lacZ* fusions cultured in HCT medium at 30°C was quantified hourly from t0 to t5. Two independent experiments are presented.

704 Figure S4. SinR is not involved in the regulation of *krs* transcription.

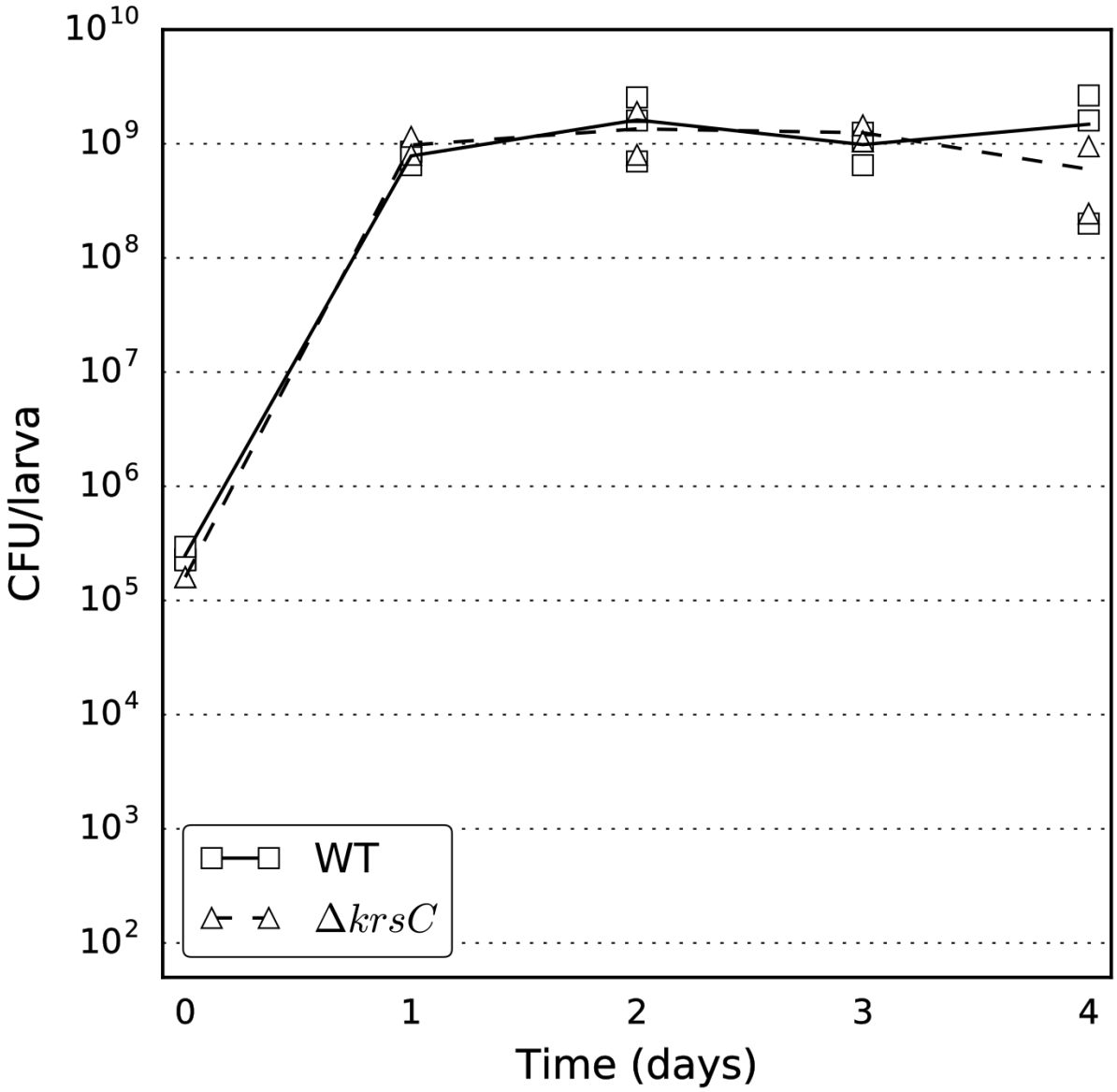
705



706 Relative β-galactosidase specific activity generated by the FL construct in the Δ*sinI-sinR* strain in

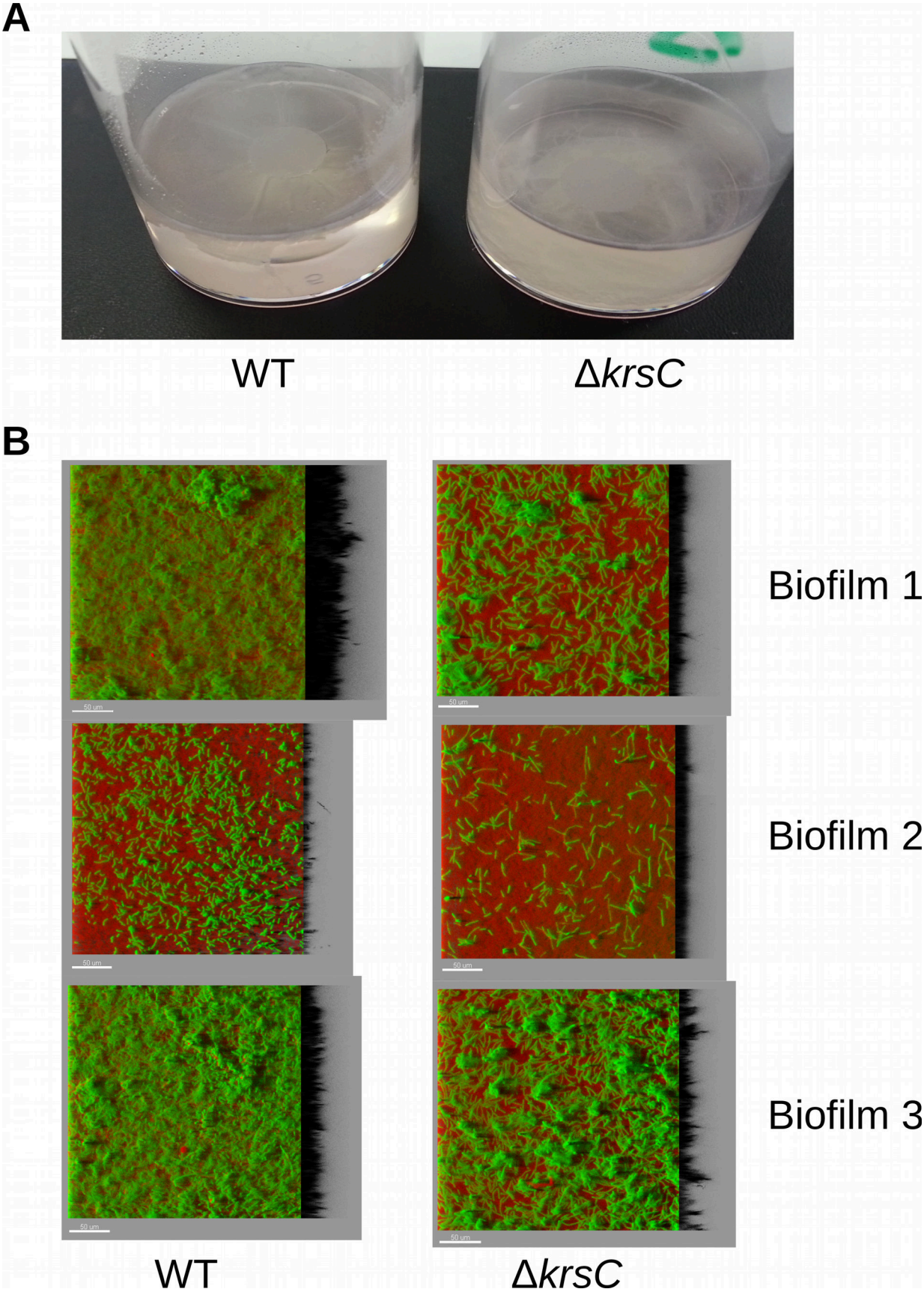
707 HCT at 30°C, normalized to its activity in the WT strain. Measures were taken hourly starting at t2.

Figure S5. The *krs* operon is not essential for necrotrophism.



CFU were counted daily from day 0 (infection) to day 4. 3 and 2 independent infections with the WT and Δ*krsC* strains, respectively, are plotted together with their respective means.

Figure S6. Biofilms formed by *B. thuringiensis* 407 Δ*krsC* are thinner and contain fewer cells than biofilms formed by the WT strain.



A. Representative macroscopic picture of biofilms formed by the WT and Δ*krsC* strains in the conditions used for confocal microscopy. **B.** 3D reconstructions from confocal microscopy of all three WT and Δ*krsC* biofilms (representative fields). Note that the second WT biofilm fell off its filter during sample preparation, before observation.

721 Movie S1. Pellicles formed by *B. thuringiensis* 407 WT and *B. thuringiensis* 407 Δ*krsC* react
722 differently to pressure applied from above.

Chapitre II - Le régulon Spo0A de *Bacillus thuringiensis* est nécessaire à son nécrotrophisme

Spo0A est le régulateur principal de la différenciation chez *B. subtilis* (González-Pastor *et al.*, 2003; Grau *et al.*, 2015; Higgins and Dworkin, 2012; Vlamakis *et al.*, 2013). Son régulon direct a été défini (Molle *et al.*, 2003a) et il a été montré que son affinité variable pour différentes boîtes Spo0A, ainsi que les nombreuses boucles de rétro-contrôle sur son expression et son activation, permettent à *B. subtilis* d'échelonner l'expression du régulon dans le temps (Fujita *et al.*, 2005). Si les régulations de Spo0A, et par Spo0A, ont été très bien décrites chez *B. subtilis*, peu de données sont disponibles dans les modèles du groupe *B. cereus*. Nous avons, dans le chapitre précédent, démontré l'implication de Spo0A dans l'expression du locus *krs*, responsable de la synthèse de la kurstakine. On sait également que Spo0A régule, directement et indirectement, la transcription de plusieurs facteurs de virulence de *B. cereus* et *B. thuringiensis*, de gènes impliqués dans la synthèse et l'accumulation de PHB, ainsi que de *plcR* (Agaisse and Lereclus, 1994; Chen *et al.*, 2010; Guillemet *et al.*, 2010; Kamar, 2014; Lereclus *et al.*, 2000; Lücking *et al.*, 2009; Poncet *et al.*, 1997). Il a de plus été montré que Spo0A est nécessaire chez *B. thuringiensis*, comme chez *B. subtilis*, à la formation de biofilm et à la sporulation (Fagerlund *et al.*, 2014; Lereclus *et al.*, 1995). Ces données laissent supposer le rôle pléiotrope de Spo0A chez *B. thuringiensis*. Nous avons donc décidé de définir le régulon Spo0A chez *B. thuringiensis*, en nous intéressant particulièrement à son implication dans le cycle infectieux de la bactérie dans *Galleria mellonella*.

1 Article 2 : « The Spo0A regulon of *Bacillus thuringiensis* is required for its necrotrophism »

1.1 *spo0A* est nécessaire au nécrotrophisme de *B. thuringiensis*

Nous avons dans un premier temps déterminé le phénotype de la souche mutante *B. thuringiensis* 407 Δ*spo0A*, par rapport à la souche parentale, lors de cultures *in vitro* et au cours de l'infection de larves de l'hôte modèle *Galleria mellonella*. Nous avons observé un léger retard de croissance de la souche mutante dans la phase exponentielle lors d'une culture

in vitro en milieu LB à 30°C (Fig. 1). Dans ces mêmes conditions, nous avons mesuré l'expression de spo0A au cours de la culture d'une souche WT (Fig. 2). Celle-ci démarre dès la fin de la phase exponentielle (t_0), avec une augmentation d'expression relativement faible entre t_0 et t_1, puis une augmentation plus forte entre t_1 et t_5.

Nous avons ensuite mesuré le temps létal (temps nécessaire à un pathogène pour tuer son hôte) de la souche mutante et observé qu'il était supérieur à celui de la souche parentale, avec une médiane à 6 h pour la WT, contre 10 h pour la Δ*spo0A* (Fig. 3). Nous avons également dénombré quotidiennement la population totale de *B. thuringiensis* de l'injection (j_0) à j_4 (Fig. 4). Cette expérience nous a permis de montrer que la souche mutante, à l'inverse de la souche parentale, est incapable de maintenir une population stable dans le cadavre de *Galleria mellonella*, autrement dit, d'être nécrotrophe. La sporulation ne semble pas être requise pour survivre sur le cadavre de l'hôte, car même 7 jours après infection on y retrouve en majorité des cellules végétatives. Il a également été montré qu'une souche *B. thuringiensis* 407 Δ*sigK*, affectée plus tard dans la voie de sporulation (et uniquement cette voie) est toujours nécrotrophe (Dubois *et al.*, 2012). Cela nous amène donc à proposer qu'une partie du régulon Spo0A, autre que la sporulation, serait nécessaire au nécrotrophisme.

1.2 Le régulon Spo0A de *B. thuringiensis*

Pour tenter de valider cette hypothèse, nous avons ensuite défini le régulon Spo0A de *B. thuringiensis*. Avant cela, nous avons effectué un travail préparatoire de ré-annotation fonctionnelle de notre génome modèle (Sheppard *et al.*, 2013). Nous avons, pour chaque gène du génome (chromosome et plasmides) tenté de définir la ou les catégorie(s) COG avec EggNOG (Fig. 5) (Powell *et al.*, 2014; Tatusov *et al.*, 1997), les données « Gene Ontology » (GO) avec InterProScan (Jones *et al.*, 2014), l'analogie avec une protéine de structure connue avec HHSearch (Söding, 2005), la présence d'un analogue chez *B. subtilis* avec BLASTp (Altschul *et al.*, 1997), et l'appartenance à plusieurs régulons déjà définis chez *B. thuringiensis*, *B. anthracis* et *B. subtilis* (Château *et al.*, 2013; Dubois *et al.*, 2012; Eijlander *et al.*, 2014; Fagerlund *et al.*, 2014; Gohar *et al.*, 2008). Cette ré-annotation (Tab. S4) nous a permis d'attribuer une fonction potentielle, plus ou moins précise, à 22 % (474/2144) des gènes dont le produit était simplement annoté « hypothetical protein ». La fonction des gènes plasmidiques reste mal annotée, avec 67 % de produits (605/906) toujours sans annotation.

En parallèle de ce travail de ré-annotation fonctionnelle, nous avons également tenté d'établir un régulon Spo0A *in silico*, par prédiction des boîtes Spo0A dans le génome de *B. thuringiensis* 407. Tout gène possédant au moins une boîte Spo0A prédite en 5' de sa phase codante a été retenu. Par simplification, les gènes contenant une ou des boîte(s) Spo0A uniquement dans leur phase codante n'ont pas été retenus, bien qu'on ne puisse pas éliminer la possibilité que ces boîtes soient fonctionnelles (Belitsky and Sonenshein, 2013).

Nous avons ensuite effectué des expériences de RNA-seq sur des cultures de *B. thuringiensis* 407 WT et Δ*spo0A* *in vitro*, en LB 30°C, à t_2 et t_5, ainsi qu'*in vivo*, dans le cadavre de *Galleria mellonella* 36 h *post mortem*. La confrontation des données *in silico*, *in vitro* et *in vivo* nous a permis de proposer un ensemble de 184 UT régulées, potentiellement directement, par Spo0A et constituant donc le régulon Spo0A central de *B. thuringiensis* (Fig. 6, Tab. S5). Le groupement de ces 184 UT par fonction COG confirme et souligne le pléiotropisme de Spo0A chez *B. thuringiensis* (Fig. 7).

Spo0A régule notamment l'expression de nombreux gènes impliqués dans le transport et le métabolisme des acides aminés, et de nombreux gènes codant pour d'autres régulateurs transcriptionnels. Nous ajoutons également plusieurs facteurs de virulence au régulon Spo0A central de *B. thuringiensis*. La régulation des gènes de sporulation se révèle n'être que partiellement conservée entre *B. thuringiensis* et *B. subtilis*. Dans le régulon central de *B. thuringiensis*, la transcription de trois UT (*spo0A*, *spoIIE* et *spoIIAA*) est activée par Spo0A, comme décrit chez *B. subtilis*. Quatre autres UT du régulon Spo0A central de *B. thuringiensis* (*ynzD*, *spoIIB*, *spoIID* et *spoIIQ*) ont un analogue chez *B. subtilis* qui n'est pas décrit comme sous le contrôle direct de Spo0A dans ce modèle. On note inversement que *spoIIGA* et *spo0F* sont absents de notre régulon central, alors qu'ils sont sous le contrôle direct de Spo0A chez *B. subtilis*.

Enfin, nous développons deux des sous-régulons potentiels de Spo0A : les gènes de mobilité et les régulons *abrB*.

1.3 Le sous-régulon potentiellement impliqué dans le nécrotrophisme

Une fois le régulon Spo0A central défini, nous avons tenté d'y rechercher les effecteurs potentiellement impliqués dans le nécrotrophisme de *B. thuringiensis*. Pour cela,

nous avons choisi de nous focaliser sur les UT du régulon central, activées ou réprimées par Spo0A et sur-exprimées *in vivo*, dans le cadavre de *Galleria mellonella* par rapport aux données *in vitro* (en LB à t_5) (Fig. 8, Tab. S6). L'analyse de ces 40 UT est en cours.

Article 2 : « The Spo0A regulon of *Bacillus thuringiensis* is required for its necrotrophism »

Manuscrit en préparation

The Spo0A regulon of *Bacillus thuringiensis* is required for its necrotrophism

Sébastien Gélis-Jeanvoine[1], Didier Lereclus[1] and Leyla Slamti[1]*

[1]Micalis Institute, INRA, AgroParisTech, Université Paris-Saclay, 78350 Jouy-en-Josas, France

*For correspondence. E-mail: leyla.slamti@jouy.inra.fr; Tel.: +33 (0) 1 34 65 23 82.

8 **Abstract**

9

Introduction

Spo0A is a well-known master transcriptional regulator that has been studied in great details in *Bacillus subtilis*. In this organism, it controls many differentiation pathways: sliding motility (Grau et al. 2015), cannibalism (González-Pastor et al. 2003), biofilm formation (Vlamakis et al. 2013), and of course sporulation (Higgins & Dworkin 2012). Phosphorylation-activated Spo0A~P activates or represses gene expression by binding to the so-called Spo0A box. Its direct regulon has been defined in *B. subtilis* and it was shown that thanks to different binding affinities for Spo0A boxes and complex feedback loops, Spo0A~P regulates gene expression in a dose-dependent manner (Molle et al. 2003; Fujita et al. 2005; Chastanet et al. 2010).

Bacillus thuringiensis (*Bt*) is another Gram positive model bacterium of the *Bacillus cereus* group. *Bt* is a well-known entomopathogen widely used as a biopesticide, and its infectious cycle inside host larvae is increasingly understood (Slamti et al. 2014; Raymond et al. 2010). The first steps of the oral infection require the PlcR quorum-sensing regulator (Salamitou et al. 2000). PlcR regulates the expression of several toxins (Nhe, Hbl, the cereolysin and the cytotoxin K) and many degradative enzymes, including phospholipases and proteases (Gohar et al. 2008). After host death, *Bt* is able to survive in its cadaver for at least 7 days (Dubois et al. 2012; Gélis-Jeanvoine unpublished observations). This behavior was named necrotrophism and shown to require the NprR quorum-sensing regulator (Dubois et al. 2012). NprR, like PlcR, regulates the expression of numerous degradative enzymes including proteases, esterases, lipases, chitinases and a chitosanase. This wide spectrum of activities is thought to allow *Bt* to use the cadaver as a source of nutrients. It was shown that a sporulation-defective mutant (Δ*sigK*) was still capable of surviving in the host cadaver, supporting the hypothesis that during necrotrophism *Bt* population remains (at least partly) vegetative (Dubois et al. 2012).

The role of Spo0A in the infectious cycle of *Bt* is often summarized as limited to the dissemination stage, allowing spore formation after a successful infection (Dubois et al. 2012;

Slamti et al. 2014). However the pleiotropism of Spo0A in several model organisms, including *Bt*, suggests that it could be involved in other parts of the life cycle of *Bt*. It was notably shown that Spo0A regulates the expression of the so-called Cry toxins (Poncet et al. 1997; Agaisse & Lereclus 1994), of the *dlt* operon involved in the D-alanylation of teichoic acids (R. Kamar & V. Sanchis, personal communication), of genes involved in the synthesis and accumulation of PHB (Chen et al. 2010), of the InhA2 metalloprotease (Guillemet et al. 2010) and indirectly, through AbrB, of the emetic toxin locus of the closely related *B. cereus* (Lücking et al. 2009). Spo0A is also known to repress the expression of the PlcR regulator, probably by direct binding to two Spo0A boxes (Lereclus et al. 2000). Two phenotypes were associated with the deletion of *spo0A* in *Bt*: the inability to form biofilm (Fagerlund et al. 2014) and of course, spores (Lereclus et al. 1995).

Here we show that Spo0A is strictly required during the necrotrophic step of the infection of *Galleria mellonella* by *Bt*. In contrast, we also show that Spo0A is involved in, but not required for, the initial infection stage. After a functional re-annotation of our reference *Bt* genome, we then define the Spo0A regulon by RNA sequencing cultures of *Bt* WT and Δ*spo0A in vitro* 2 and 5 hours after the end of exponential phase and *in vivo* in the host cadaver 36 hours *post mortem*. Comparing RNA-seq datasets with *in silico* genome-wide annotation of putative Spo0A boxes, we propose a "core regulon" of 184 potentially directly Spo0A-regulated genes in *Bt*. We finally focus on a smaller subset of 40 Spo0A-regulated genes that are likely to be more directly involved in the necrotrophism of *Bt*.

Experimental procedures

RNA extraction from shaken liquid cultures

2 mL aliquots from LB cultures were sampled 2 and 5 hours after the end of exponential phase (t_2 and t_5). Samples were mixed with an equal volume of RNA*later* (Life Technologies, Ambion, USA) and stored at -20°C until RNA extraction. Total RNA was then prepared as described previously for RT-PCR (Gélis-Jeanvoine *et al*., 2016). RNA Integrity Numbers (RINs) measured with a Nano chip on a BioAnalyzer 2100 (Agilent Technologies) ranged between 8.5 and 9.9, indicating high-quality extractions (Schroeder et al. 2006).

Infection of axenic Galleria mellonella *with* Bacillus thuringiensis *str. 407*

In vivo experiments were adapted from Salamitou *et. al.* (Salamitou et al. 2000). *G. mellonella* eggs were irradiated in a laminar flow hood by a UV lamp for 5 minutes. Eggs were then hatched at 30°C in a sterile box containing γ-sterilized beeswax and pollen. Despite our best efforts, larvae still contained approximately 500 CFUs of *Enterococcus mundtii*, their most represented symbiont (C. Nielsen-Leroux, personal communication). Larvae were then infected by intrahaemocœlic injections at the base of the last prolegs with 200.000 vegetative cells in exponential phase. Infected larvae were incubated at 30°C under a camera. After insects death, movies were analyzed to determine the exact time of death for each larva. A larva was considered dead when it stopped moving (melanization occurred roughly 2 hours later). Inocula were numerated by plating and counting CFUs after injections. When mentioned, total CFUs were numerated in larvae cadavers by grinding 2 cadavers in 2mL of physiological water, adjusting the level to approximately 10mL with physiological water and plating 10 µL on LB plates. Raw CFU numerations are then expressed as the number of CFUs per larva cadaver.

RNA extraction from dead larvae

Bacteria were harvested from larvae 20 hours and 36 hours *post mortem*. Bacteria isolation from cadavers was adapted from (Verplaetse et al. 2015). Larvae were quickly cut open, immersed in RNA*later* (Life Technologies, Ambion, USA) and vortexed. Tubes were stored in a fridge overnight to allow thorough penetration of RNA*later*. The soluble fraction was then transferred to a new tube and centrifuged 10 minutes at 14.000 g at 4°C. The fat pellicle and supernatant were discarded and the pellet resuspended in 500µL of 0.22 µm filtered PBS. The suspension was then filtered on cotton in a 1 mL syringe to remove larvae debris, and the flow-through was mixed with an equal volume of RNA*later*. Tubes were kept overnight at 4°C and stored at -20°C prior to RNA extraction. RINs of total RNA extracts from larvae cadavers ranged between 7 and 8.5. The lower quality compared to RNA extracted from LB cultures was expected, and the measured RINs are still compatible with further RNA-seq protocols.

cDNA library generation and sequencing

rRNA was depleted from total RNA extracts with the Ribo-Zero kit (Epicentre, Wisconsin, USA), following manufacturer's instructions with an input of 4.5 µg of total RNA. For some total RNA samples from insect cadavers a second round of Ribo-Zero depletion was performed because BioAnalyzer spectra showed residual contamination with rRNA. Stranded Illumina cDNA libraries were then generated from rRNA-depleted RNA extracts using the ScriptSeq kit for bacteria (Epicentre, Wisconsin, USA), following manufacturer's instructions. During the process, rRNA-depleted samples were purified with the RNeasy MinElute Cleanup kit (Qiagen, France) and both the cDNA and final libraries were purified with the MinElute PCR purification kit (Qiagen, France). Libraries were sequenced in an Illumina HiSeq1000 sequencer with the TruSeq SBS kit v3 – HS 50 cycles (Illumina). Sequencing depths ranged from 11 to 26 million reads per library.

RNA-seq workflow

Reads were demultiplexed with CASAVA (v1.8.2) (Illumina) and trimmed of their adapters with Cutadapt (v1.3, parameters "-n 5 -O 4 -m 10") (Martin 2011). Quality of the raw and trimmed reads was assessed with FastQC (v0.10.1) (Babraham Bioinformatics). Reads from the larvae cadavers libraries were subjected to taxonomic assignment with Kraken (v 0.10.5) (Wood & Salzberg 2014), using the MiniKraken database (2014-12-08 update), to assess the relative presence of microorganisms other than *Bt*. Reads were then mapped on the *Bacillus thuringiensis* str. 407 reference genome (chromosome: INSDC CP003889.1, plasmids: INSDC CP003890.1 to CP003898.1) with BWA (v0.7.10) (Li & Durbin 2009), allowing a maximum edit distance of 0.04. The resulting SAM files were filtered to keep only single-mapping reads (marked with the "XT:A:U" flag). Mappings were visualized with Tablet (Milne et al. 2013) to ensure reliability of the mapping. Reads per gene were numerated with the HTSeq-count script (v0.5.4p3, parameters "--stranded=yes --a=10 --type=gene --idattr=Name") (Anders et al. 2014). Differentially expressed genes (DEGs) were determined with R (v3.1.1) (R Core Team 2016), Bioconductor (v2.28.0) (Gentleman et al. 2004) and DESeq2 (v1.8.1) (Love et al. 2014). DEGs with $p_{adj} \leq 10^{-2}$ and $-1 \geq \log_2$ ratio ≥ 1 were used in further analysis.

Functional re-annotation of the reference Bt *407 genome*

All the genes of our reference genome were functionally re-annotated using InterProScan (Jones et al. 2014), HHSearch (Söding 2005) and EggNOG (Powell et al. 2014). InterProScan v. 5 was queried online through its SOAP API (http://www.ebi.ac.uk/Tools/webservices/services/pfa/iprscan5_soap) with the --goterms and --nopathways flags. GO terms were then parsed and used as three different annotations (process, function and component). HHSearch was used on our local cluster to query a HMM database prepared from the PDB70 database (2014-09-06 update). To assign COG letters to our query genes, we first aligned them with BLASTp (e value threshold of 10^{-2}) against the EggNOG v. 4.0 database

(http://eggnogdb.embl.de/download/eggnog_4.0/eggnogv4.proteins.core_periphery.fa.gz). The best hit for each query gene was then used to retrieve the cognate COG letter *via* an in-house script (EggnogGenome, https://github.com/seb-ksl/EggnogGenome). Different published regulons were also used as data sources to further enrich our annotation: the Spo0A and σ^F regulons of *Bacillus subtilis* (from the SporeWeb server (Eijlander et al. 2014)), the PlcR (Gohar et al. 2008), NprR (Dubois et al. 2012) and SinR (Fagerlund et al. 2014) regulons of *Bt* and the CodY regulon of *Bacillus anthracis* (Château et al. 2013). Putative homologues in the *B. subtilis* 168 genome (BLASTp e value threshold of 10^{-2}) were also annotated.

In silico prediction of the Spo0A boxes in the Bt *407 genome*

Putative Spo0A boxes were searched in the genome (chromosome and plasmids) of *Bt* 407 with the Virtual Footprint tool, using the "Spo0A II" motif of *B. subtilis* 168 (PRODORIC accession MX000019) as a query (Münch et al. 2005). Boxes identified with this program were then filtered to retain only those in the 5' region of genes, using our reference annotation. We chose to discard putative boxes that fell in annotated genes for simplicity, even though such boxes were shown to be sometimes functional (Belitsky & Sonenshein 2013). The Spo0A box defined in *B. subtilis* is very likely to be conserved in *Bt* because of the high degree of conservation of Spo0A of *Bt* and Spo0A of *B. subtilis* (81% identity, 89% similarity) and especially of its C-terminal DNA-binding domain (Rowe-Magnus & Spiegelman 1998) (92% identity, 97% similarity). Experimentally, it was also shown that the two *in silico*-detected Spo0A boxes upstream of *plcR* were actually protected of DNase I degradation when incubated with Spo0A~P of *B. subtilis* (Lereclus et al. 2000).

Results

*Behavior of the Δ*spo0A *mutant* in vitro *and* in vivo

The *Bacillus thuringiensis* 407 (*Bt* 407) Δ*spo0A* mutant exhibited slightly reduced division rate during exponential phase in LB at 30°C compared to the WT strain (Fig. 1). However growth of the mutant was less impaired at 30°C than was previously reported at 37°C (Fagerlund et al. 2014). Monitoring the transcription of a P_{spo0A}-*lacZ* fusion from a plasmid showed that a low expression was already detected at the end of exponential phase (t_0) and that it steadily increased between t_1 and t_5 (Fig. 2).

The mutant strain was less efficient in killing larvae of *Galleria mellonella*. While $2 \cdot 10^5$ CFU of the WT strain killed 100% (63/63) of the infected larvae, the same amount of Δ*spo0A* strain killed 89% (57/63) of infected larvae. Infections with the Δ*spo0A* mutant killed the larvae more slowly than infections with the parental strain (Fig. 3). Additionally to being able to kill larvae, it was shown previously that *Bt* is able to maintain a vegetative population in the cadaver of the host it killed, a behavior called necrotrophism (Dubois et al. 2012). We wanted to test a possible involvement of *spo0A* in necrotrophism and followed the number of CFUs after infections with the WT and Δ*spo0A* strains (Fig. 4). This allowed us to show that *spo0A* is indeed required for *Bt* to become necrotrophic, because unlike the WT strain, the mutant strain was unable to maintain a stable population after initial multiplication. Previous experiments already showed that sporulation *per se* is not a requirement of necrotrophism (Dubois et al. 2012), suggesting that other Spo0A-dependent pathways could be involved in the process. However there is, to our knowledge, no report to date of the Spo0A regulon in any member of the *Bacillus cereus* group.

Functional re-annotation of the reference Bt *407 genome*

We used the publicly available complete annotated genome of *Bt* 407 Cry^- as our reference genome (Sheppard et al. 2013). To enrich the initial annotation, we performed full genome

InterProScan, HHSearch and EggNOG annotations with in-house scripts (see experimental procedures, Table S4) (Jones et al. 2014; Söding 2005; Powell et al. 2014). The EggNOG annotation confirmed previous reports, with a high representation of genes involved in amino acid transport and metabolism (Fig. 5) (Ivanova et al. 2003). Additionally to enriching existing annotations, this approach allowed us to propose putative annotations for 22% (474/2144) of genes whose predicted product was only annotated as “hypothetical protein”. However, most genes carried by the 9 plasmids remained fully hypothetical: 67% (605/906) of all the plasmidic genes had no initial annotation (“hypothetical protein”) and still have no putative homologue in *B. subtilis* 168, no COG letter (or an S), no GO term and no HHSearch prediction.

In silico *prediction of the Spo0A regulon*

Putative Spo0A boxes were searched in the genome (chromosome and plasmids) of *Bt* 407. 7464 boxes were found and were then filtered to retain only the 729 boxes that were in the 5’ region of annotated genes. Because 5’ regions can be shared between two head-to-tail genes, these 729 boxes were attributed to 828 genes. Genes belonging to this virtual regulon had, most frequently, a single predicted Spo0A box in their 5’ region, but some genes had up to 5 predicted boxes. A position-weight matrix of the 729 selected putative boxes was generated and was very similar to the motif defined in *Bacillus subtilis*, showing no enrichment of a specific motif in the *Bt* boxes (data not shown).

*RNA-seq analysis of the WT and Δ*spo0A *transcriptomes*

Total RNA was extracted from shaken LB 2 and 5 hours after the end of exponential phase (t_2 and t_5) and 36 hours *post mortem* from the *Galleria mellonella* larvae cadavers. Total RNA was then transformed into directional cDNA Illumina library and sequenced in a HiSeq 1000 (Illumina). Reads from the *in vivo* libraries were subjected to taxonomic assignment to assess purity of the

libraries. In all four libraries, reads that could be assigned fell in vast majority in the *Bt* lineage, from *Firmicutes* to *Bt* itself. Some reads, however, revealed the presence of other organisms including *Enterobacteriaceae*, an *Enterobacteria* phage, and a *Lepidoptera* virus. 94% to 98% of reads from the LB libraries could be mapped to our *Bt* 407 reference genome, and 67% to 96% for larvae libraries. After filtering out multiple-mapping reads, the smallest mapping still contained 10.5 million exploitable reads, a number compatible with further analyses of differential gene expression (Haas et al. 2012). Final mappings were all visualized with Tablet to assess their reliability (Milne et al. 2013). Mapping coordinates and strands were, in most observed cases, in agreement with CDSs from the reference chromosome annotation. However reads from the plasmids often showed poor correlation with existing annotations, reflecting the limited reliability of *in silico* CDSs annotation of totally unknown genes with no homologue. Disruption of the *spo0A* gene with the kanamycin resistance cassette (Lereclus et al. 1995) was visible in mappings of this strain, confirming that no cross-contamination between WT and mutant cultures occurred. Raw counts of reads per gene was performed with the HTseq-count script (Anders et al. 2014) and differential gene expression was subsequently assessed with the DESeq 2 R package (Love et al. 2014; R Core Team 2016). Reliability of the DESeq 2 analyses was assessed by plotting PCA and dendrograms of sample to sample distances. Both representations confirmed that variances between biological replicates were negligible towards variances between WT and Δ*spo0A* strains (data not shown).

The core Spo0A regulon of Bt

1987 genes were found to be at least two-fold differentially expressed between the WT and Δ*spo0A* strains at t_2 in LB and 2534 at t_5, forming a total set of 3388 *in vitro spo0A*-dependent genes. Of these 3388 genes, 41% (1402) were differentially expressed only at t_5, 25% (860) only at t_2, 18% (607) were down-regulated in the Δ*spo0A* mutant at both time points, 12% (417) were up-

regulated at both time points and 3% (102) were up-regulated at t_2 and down-regulated at t_5 or the opposite. In cadavers of *Galleria mellonella* larvae 36 hours *post mortem*, the expression of 2034 genes was found to be *spo0A*-dependent. Overlap of the *in vitro*, *in vivo* and *in silico* genes sets allow to define the core Spo0A regulon: 184 genes that harbor at least one Spo0A box in their 5' upstream region and were found to be *spo0A*-dependent both *in vitro* and *in vivo* (Fig. 6).

A global functional analysis of the core regulon highlights the pleiotropism of Spo0A (Fig. 7, Table S5). Spo0A regulates, potentially directly *via* Spo0A boxes, the transcription of genes involved in most functional categories. The most represented function in the core regulon is "amino acid transport and metabolism" (category "E") which, as highlighted earlier and in previous studies, is thought to play an important role in the metabolism of *Bacillus cereus sensu lato* (Ivanova et al. 2003). On the 12 genes that we found to be potentially directly repressed by Spo0A, 9 are annotated as involved in the import of amino acids (including *dppE4*, the first gene of the *opp* operon), *vs.* only 1 among the 6 Spo0A-activated genes. The overall involvement of Spo0A in amino acid metabolism could then be the repression of amino acid import. This switch could notably lead, besides a drastic change in metabolic pathways, to a drop in PlcR activity because the Opp machinery is involved in the import of the mature signaling peptide PapR (Gominet et al. 2001).

Strikingly, Spo0A was found to regulate potentially directly the expression of 17 transcriptional regulators (category "K"), 6 through activation and 11 through repression. These 17 regulators exclude those involved in sporulation, thus Spo0A is not counted. Among the Spo0A-repressed regulators-encoding genes are *abrB1* and *plcR*, confirming previous reports in *B. subtilis* and *Bt* (Lereclus et al. 2000; Strauch et al. 1990). The *glcT* gene, predicted to encode GlcT, the *ptsGHI* antiterminator (Stülke et al. 1997), was also found to be repressed by Spo0A. Among the Spo0A-activated genes are *mogR* and *treR*, predicted to encode respectively, a motility repressor, and a trehalose operon repressor. The 12 other regulators have elusive annotations and no putative function could be attributed to them as of today. However, the abundance of transcriptional

regulators under the potential direct regulation of Spo0A can account for the high number of differentially expressed genes between the WT and Δ*spo0A* strains. The most striking absence in this category, compared to *B. subtilis*, is SinI, the anti-repressor of SinR (Bai et al. 1993; Shafikhani et al. 2002). Indeed, *sinI* was found to be down-regulated in the *Bt* 407 Δ*spo0A* strain compared to the WT only in LB at t_2, and most importantly, has no predicted Spo0A box in its 5' region.

Spo0A is also increasingly recognized as a regulator of virulence factors. It was notably shown to repress in *Bt*, both directly and indirectly, the transcription of Cry toxins (Poncet et al. 1997; Agaisse & Lereclus 1994), of the InhA2 metalloprotease (Guillemet et al. 2010) and indirectly, *via* AbrB, of the emetic toxin of the closely related *Bacillus cereus* (Lücking et al. 2009). Our model strain lacks *ces* and *cry* plasmidic genes, but we confirm that *inhA2* (*ina1*) belongs to the core Spo0A regulon and add *cytK* (*hlgB*), *nprB* and *nprM* and *nprS* to the lengthening list of Spo0A-regulated virulence factors. *cytK* was remarkably strongly repressed by Spo0A in all conditions, with measured ratios of 20 (larvae cadavers), 178 (LB t_2) and 218 (LB t_5). This strong repression can be explained by the presence of two close canonical Spo0A boxes downstream of its transcription start (Brillard & Lereclus 2004). *nprM* and *nprS* were found to be repressed by ratios between 11 and 29 in all conditions, while *nprB* was activated with smaller ratios (5 and 8, not detected in LB at t_5). Interestingly, *nprM* and *nprS*, both repressed by Spo0A, also belong to the PlcR regulon, while *nprB*, Spo0A-activated, belongs to the NprR regulon. The potential direct regulation of *nprS* by Spo0A remains unclear because its two putative Spo0A boxes are shared with *plcR*, as the two genes are in head-to-tail orientation.

Sporulation genes

As expected, Spo0A was found to regulate the expression of several genes involved in sporulation (Fig. 7, category "X"). It activates the transcription of 9 genes, among which only *spo0A*, *spoIIE* and *spoIIAA* (in putative operon with *sigF2*) are also activated directly by Spo0A in

B. subtilis (Molle et al. 2003). Of the 6 remaining Spo0A-activated genes, 4 are closely related to genes known to be involved in sporulation but not directly regulated by Spo0A in *B. subtilis* (*ynzD*, *spoIIB*, *spoIID* and *spoIIQ*), one is predicted to encode a small acid-soluble protein (*sasP1*) and one is predicted to encode a protein involved in germination (*yndE1*). The only potentially directly Spo0A-repressed sporulation gene is predicted to encode another germination protein (*gerA*). Surprisingly, *spo0F* and *spoIIGA-sigE* are absent of the Spo0A core regulon of *Bt*. They were actually found to be differentially expressed between the WT and Δ*spo0A* strains both *in vitro* and *in vivo*, but the lack of a clear Spo0A box in their promoter sets them in the indirect Spo0A regulon. A Spo0A box could actually be found upstream of *spoIIGA* by forcing the algorithm, but its statistical score was too low to retain it in the final dataset. In *B. subtilis* the *spoIIGA* promoter contains 4 cryptic boxes (of which only one is detected by the *in silico* tool we used) that are nevertheless functional (Baldus et al. 1994), so we cannot formally exclude the possibility of a direct regulation by Spo0A in *Bt*.

The indirect Spo0A regulon of Bt

As mentioned earlier, the core regulon contains 17 transcriptional regulators. Among them, only PlcR has been studied in *Bt* and had its regulon defined (Gohar et al. 2008). It is therefore not surprising to see that most PlcR-regulated genes (92%, 80/87) were found to be differentially expressed in the Δ*spo0A* strain compared to the WT *in vitro* and/or *in vivo*. Some of these genes belonged to the core regulon, suggesting a possible direct regulation by both PlcR and Spo0A.

The gene encoding the putative motility repressor MogR is part of the core regulon too. In *Listeria monocytogenes* MogR was shown to repress the expression of genes involved in motility (Shen & Higgins 2006), and this function is thought to be conserved in *Bacillus cereus sensu lato* (Shen et al. 2009). No motility-related gene belonged to our core regulon, however 92% (35/38) of them were found to be differentially expressed in the Δ*spo0A* mutant in at least one condition.

Strikingly, all 35 were up-regulated in the mutant, revealing an extensive indirect repression of Spo0A towards motility genes. *mogR* was down-regulated in all tested conditions with ratios of 12 (LB t_2), 4 (LB t_5) and 22 (*in vivo*). The repression of motility genes could then be the result of a Spo0A → MogR –| motility chain of events.

Bt 407 has 4 putative paralogues of *abrB*: *abrB1* to *abrB4*. *abrB1* is the closest from the *B. subtilis abrB* (83% identity, 91% similarity) and they share the same genetic environment. *abrB2*, *abrB3* and *abrB4* are still highly similar to *abrB* of *B. subtilis* (84%, 77% and 77% similarity, respectively) but lay in different genetic contexts. All 4 genes were found to be regulated by Spo0A *in vitro* at both times, but only *abrB1* and *abrB3* were detected as differentially expressed between the WT and Δ*spo0A* strains *in vivo*. Interestingly, *abrB1* and *abrB4* were found to be repressed by Spo0A while *abrB2* and *abrB3* were Spo0A-activated.

The in vivo specific Spo0A regulon of Bt

Knowing that Spo0A is required for *Bt* to become necrotrophic, we wanted to assess which part of its regulon could be responsible for this phenotype. To do so, we computed the intersection between the core regulon and the set of genes differentially expressed between the *in vivo* and *in vitro* (LB t_5) conditions. This intersection contained 87 genes, 40 of which were up-regulated *in vivo* (Table S6, Fig. 8). Among these 40 genes, 19 were activated by Spo0A *in vivo* and 21 were repressed.

Talk about highest ratios?

Plasmids? pBTB_9p down in Δ0A (but still somewhat sequenced) => replication issue?

Discussion

Discuss why no motility phenotype in Gohar but mogR in this study: motility assay on plate involves log phase cells, when no Spo0A. In this case, motility allowed. Would require to assess motility at stat phase times in WT and Δ0A (fresh microscopy).

Discuss lack of sinI in regulon.

Conclusions on plasmids would require first re-annotation of CDSs.

References

Agaisse, H. & Lereclus, D., 1994. Expression in Bacillus subtilis of the Bacillus thuringiensis cryIIIA toxin gene is not dependent on a sporulation-specific sigma factor and is increased in a spo0A mutant. *Journal of bacteriology*, 176(15), pp.4734–41. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=196296&tool=pmcentrez&rendertype=abstract [Accessed November 5, 2015].

Anders, S., Pyl, P.T. & Huber, W., 2014. HTSeq - A Python framework to work with high-throughput sequencing data. *Bioinformatics (Oxford, England)*, 31(2), pp.166–169. Available at: http://bioinformatics.oxfordjournals.org/content/31/2/166 [Accessed September 29, 2014].

Bai, U., Mandic-Mulec, I. & Smith, I., 1993. SinI modulates the activity of SinR, a developmental switch protein of Bacillus subtilis, by protein-protein interaction. *Genes & development*, 7(1), pp.139–48. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8422983 [Accessed July 13, 2012].

Baldus, J.M. et al., 1994. Phosphorylation of Bacillus subtilis transcription factor Spo0A stimulates transcription from the spoIIG promoter by enhancing binding to weak 0A boxes. *Journal of bacteriology*, 176(2), pp.296–306. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=205050&tool=pmcentrez&rendertype=abstract [Accessed November 4, 2015].

Belitsky, B.R. & Sonenshein, A.L., 2013. Genome-wide identification of Bacillus subtilis CodY-binding sites at single-nucleotide resolution. *Proceedings of the National Academy of Sciences of the United States of America*, 110(17), pp.7026–31. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3637721&tool=pmcentrez&rendertype=abstract [Accessed November 28, 2013].

Brillard, J. & Lereclus, D., 2004. Comparison of cytotoxin cytK promoters from Bacillus cereus strain ATCC 14579 and from a B. cereus food-poisoning strain. *Microbiology (Reading, England)*, 150(Pt 8), pp.2699–705. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15289566 [Accessed July 3, 2015].

Chastanet, A. et al., 2010. Broadly heterogeneous activation of the master regulator for sporulation in Bacillus subtilis. *Proceedings of the National Academy of Sciences of the United States of America*, 107(18), pp.8486–91. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2889527&tool=pmcentrez&rendertype=abstract [Accessed July 13, 2012].

Château, A. et al., 2013. Identification of CodY targets in Bacillus anthracis by genome-wide in vitro binding analysis. *Journal of bacteriology*, 195(6), pp.1204–13. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3591999&tool=pmcentrez&rendertype=abstract [Accessed March 23, 2015].

Chen, H.-J. et al., 2010. The master transcription factor Spo0A is required for poly(3-hydroxybutyrate) (PHB) accumulation and expression of genes involved in PHB biosynthesis

in Bacillus thuringiensis. *FEMS microbiology letters*, 304(1), pp.74–81. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20100285 [Accessed February 21, 2013].

Dubois, T. et al., 2012. Necrotrophism is a quorum-sensing-regulated lifestyle in Bacillus thuringiensis. *PLoS pathogens*, 8(4), p.e1002629. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3325205&tool=pmcentrez&rendertype=abstract [Accessed July 13, 2012].

Eijlander, R.T. et al., 2014. SporeWeb: an interactive journey through the complete sporulation cycle of Bacillus subtilis. *Nucleic acids research*, 42(Database issue), pp.D685–91. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3964945&tool=pmcentrez&rendertype=abstract [Accessed November 25, 2015].

Fagerlund, A. et al., 2014. SinR Controls Enterotoxin Expression in Bacillus thuringiensis Biofilms T. M. Koehler, ed. *PLoS ONE*, 9(1), p.e87532. Available at: http://dx.plos.org/10.1371/journal.pone.0087532 [Accessed February 7, 2014].

Fujita, M., González-Pastor, J.E. & Losick, R., 2005. High- and low-threshold genes in the Spo0A regulon of Bacillus subtilis. *Journal of bacteriology*, 187(4), pp.1357–68. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=545642&tool=pmcentrez&rendertype=abstract [Accessed November 5, 2012].

Gentleman, R.C. et al., 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome biology*, 5(10), p.R80. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=545600&tool=pmcentrez&rendertype=abstract [Accessed July 11, 2014].

Gohar, M. et al., 2008. The PlcR virulence regulon of Bacillus cereus. *PloS one*, 3(7), p.e2793. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2464732&tool=pmcentrez&rendertype=abstract [Accessed July 13, 2012].

Gominet, M. et al., 2001. Oligopeptide permease is required for expression of the Bacillus thuringiensis plcR regulon and for virulence. *Molecular microbiology*, 40(4), pp.963–75. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11401703 [Accessed July 13, 2012].

González-Pastor, J.E., Hobbs, E.C. & Losick, R., 2003. Cannibalism by sporulating bacteria. *Science (New York, N.Y.)*, 301(5632), pp.510–3. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12817086 [Accessed July 9, 2015].

Grau, R.R. et al., 2015. A Duo of Potassium-Responsive Histidine Kinases Govern the Multicellular Destiny of Bacillus subtilis. *mBio*, 6(4). Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4495169&tool=pmcentrez&rendertype=abstract [Accessed September 3, 2015].

Guillemet, E. et al., 2010. The InhA metalloproteases of Bacillus cereus contribute concomitantly to virulence. *Journal of bacteriology*, 192(1), pp.286–94. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2798277&tool=pmcentrez&rendertype=abstract [Accessed November 25, 2015].

Haas, B.J. et al., 2012. How deep is deep enough for RNA-Seq profiling of bacterial transcriptomes? *BMC genomics*, 13(1), p.734. Available at: http://www.biomedcentral.com/1471-2164/13/734 [Accessed April 3, 2015].

Higgins, D. & Dworkin, J., 2012. Recent progress in Bacillus subtilis sporulation. *FEMS microbiology reviews*, 36(1), pp.131–48. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22091839 [Accessed November 9, 2012].

Ivanova, N. et al., 2003. Genome sequence of Bacillus cereus and comparative analysis with Bacillus anthracis. *Nature*, 423(6935), pp.87–91. Available at: http://dx.doi.org/10.1038/nature01582 [Accessed January 22, 2014].

Jones, P. et al., 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics (Oxford, England)*, 30(9), pp.1236–40. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3998142&tool=pmcentrez&rendertype=abstract [Accessed July 13, 2014].

Lereclus, D. et al., 1995. Overproduction of Encapsulated Insecticidal Crystal Proteins in a Bacillus thuringiensis spoOA Mutant. *Bio/Technology*, 13(1), pp.67–71. Available at: http://europepmc.org/abstract/MED/9634751 [Accessed May 7, 2015].

Lereclus, D. et al., 2000. Regulation of toxin and virulence gene transcription in Bacillus thuringiensis. *International Journal of Medical Microbiology*, 290(4), pp.295–299. Available at: http://www.sciencedirect.com/science/article/pii/S1438422100800247 [Accessed October 24, 2013].

Li, H. & Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)*, 25(14), pp.1754–60. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2705234&tool=pmcentrez&rendertype=abstract [Accessed July 9, 2014].

Love, M.I., Huber, W. & Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, 15(12), p.550. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25516281 [Accessed December 23, 2014].

Lücking, G. et al., 2009. Cereulide synthesis in emetic Bacillus cereus is controlled by the transition state regulator AbrB, but not by the virulence regulator PlcR. *Microbiology (Reading, England)*, 155(Pt 3), pp.922–31. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19246763 [Accessed October 7, 2015].

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17(1), p.10. Available at: http://journal.embnet.org/index.php/embnetjournal/article/view/200/479 [Accessed January 12, 2015].

Milne, I. et al., 2013. Using Tablet for visual exploration of second-generation sequencing data. *Briefings in bioinformatics*, 14(2), pp.193–202. Available at: http://bib.oxfordjournals.org/content/14/2/193 [Accessed March 24, 2015].

443 Molle, V. et al., 2003. The Spo0A regulon of Bacillus subtilis. *Molecular microbiology*, 50(5),
444 pp.1683–701. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14651647 [Accessed
445 February 22, 2013].

446 Münch, R. et al., 2005. Virtual Footprint and PRODORIC: an integrative framework for regulon
447 prediction in prokaryotes. *Bioinformatics (Oxford, England)*, 21(22), pp.4187–9. Available at:
448 http://www.ncbi.nlm.nih.gov/pubmed/16109747 [Accessed October 30, 2015].

449 Poncet, S. et al., 1997. Spo0A represses transcription of the cry toxin genes in Bacillus
450 thuringiensis. *Microbiology (Reading, England)*, 143 (Pt 8, pp.2743–51. Available at:
451 http://www.ncbi.nlm.nih.gov/pubmed/9274027 [Accessed February 21, 2013].

452 Powell, S. et al., 2014. eggNOG v4.0: nested orthology inference across 3686 organisms. *Nucleic*
453 *acids research*, 42(Database issue), pp.D231–9. Available at:
454 http://nar.oxfordjournals.org/content/42/D1/D231 [Accessed February 17, 2015].

455 R Core Team, 2016. R: A Language and Environment for Statistical Computing. Available at:
456 http://www.r-project.org/.

457 Raymond, B. et al., 2010. Bacillus thuringiensis: an impotent pathogen? *Trends in microbiology*,
458 18(5), pp.189–94. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20338765 [Accessed
459 September 10, 2015].

460 Rowe-Magnus, D.A. & Spiegelman, G.B., 1998. DNA strand separation during activation of a
461 developmental promoter by the Bacillus subtilis response regulator Spo0A. *Proceedings of the*
462 *National Academy of Sciences of the United States of America*, 95(9), pp.5305–10. Available
463 at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?
464 artid=20256&tool=pmcentrez&rendertype=abstract [Accessed December 3, 2015].

465 Salamitou, S. et al., 2000. The plcR regulon is involved in the opportunistic properties of Bacillus
466 thuringiensis and Bacillus cereus in mice and insects. *Microbiology*, 146(11), pp.2825–2832.
467 Available at: http://mic.sgmjournals.org/content/146/11/2825.long [Accessed January 29,
468 2013].

469 Schroeder, A. et al., 2006. The RIN: an RNA integrity number for assigning integrity values to RNA
470 measurements. *BMC molecular biology*, 7, p.3. Available at:
471 http://www.pubmedcentral.nih.gov/articlerender.fcgi?
472 artid=1413964&tool=pmcentrez&rendertype=abstract [Accessed July 10, 2014].

473 Shafikhani, S.H. et al., 2002. Postexponential regulation of sin operon expression in Bacillus
474 subtilis. *Journal of bacteriology*, 184(2), pp.564–71. Available at:
475 http://www.pubmedcentral.nih.gov/articlerender.fcgi?
476 artid=139554&tool=pmcentrez&rendertype=abstract [Accessed October 16, 2015].

477 Shen, A. & Higgins, D.E., 2006. The MogR transcriptional repressor regulates nonhierarchal
478 expression of flagellar motility genes and virulence in Listeria monocytogenes. *PLoS*
479 *pathogens*, 2(4), p.e30. Available at: http://journals.plos.org/plospathogens/article?
480 id=10.1371/journal.ppat.0020030 [Accessed November 30, 2015].

Shen, A., Higgins, D.E. & Panne, D., 2009. Recognition of AT-rich DNA binding sites by the MogR repressor. *Structure (London, England : 1993)*, 17(5), pp.769–77. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2712671&tool=pmcentrez&rendertype=abstract [Accessed November 30, 2015].

Sheppard, A.E. et al., 2013. Complete Genome Sequence of Bacillus thuringiensis Strain 407 Cry-. *Genome announcements*, 1(1). Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3569317&tool=pmcentrez&rendertype=abstract [Accessed September 17, 2015].

Slamti, L. et al., 2014. Quorum sensing in Bacillus thuringiensis is required for completion of a full infectious cycle in the insect. *Toxins*, 6(8), pp.2239–55. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4147580&tool=pmcentrez&rendertype=abstract [Accessed September 28, 2015].

Söding, J., 2005. Protein homology detection by HMM-HMM comparison. *Bioinformatics (Oxford, England)*, 21(7), pp.951–60. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15531603 [Accessed May 18, 2015].

Strauch, M. et al., 1990. The SpoOA protein of Bacillus subtilis is a repressor of the abrB gene. *Proceedings of the National Academy of Sciences of the United States of America*, 87(5), pp.1801–5. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=53571&tool=pmcentrez&rendertype=abstract [Accessed October 7, 2015].

Stülke, J. et al., 1997. Induction of the Bacillus subtilis ptsGHI operon by glucose is controlled by a novel antiterminator, GlcT. *Molecular microbiology*, 25(1), pp.65–78. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11902727 [Accessed November 25, 2015].

Verplaetse, E. et al., 2015. Cell Differentiation in a Bacillus thuringiensis Population during Planktonic Growth, Biofilm Formation, and Host Infection. *mBio*, 6(3), pp.e00138–15. Available at: http://mbio.asm.org/lookup/doi/10.1128/mBio.00138-15.

Vlamakis, H. et al., 2013. Sticking together: building a biofilm the Bacillus subtilis way. *Nature reviews. Microbiology*, 11(3), pp.157–68. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23353768 [Accessed May 24, 2013].

Wood, D.E. & Salzberg, S.L., 2014. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome biology*, 15(3), p.R46. Available at: http://genomebiology.com/2014/15/3/R46 [Accessed November 24, 2014].

Acknowledgments

We are indebted to Christophe Buisson for rearing the insects. We thank Michel Gohar, Stéphane Perchat and Cyprien Guérin for helpful discussions. This work has benefited from the facilities and expertise of the high throughput sequencing platform of IMAGIF (Centre de Recherche de Gif – http://www.imagif.cnrs.fr). We are grateful to the INRA MIGALE bioinformatics platform (http://migale.jouy.inra.fr) for providing computational resources.

Legends to figures

Figure 1: The Δ*spo0A* strain is slightly impaired in vegetative growth in shaken LB at 30°C.

Optical density of shaken LB cultures at 30°C were taken from inoculation to t_5 (5 hours after the end of exponential phase). In black circles (●), the *Bt* 407 WT strain. In white circles (○), the *Bt* 407 Δ*spo0A* strain. Measures of three independent experiments are plotted.

Figure 2: Expression of *spo0A* during the stationary phase.

β-galactosidase specific activity (expressed in U/mg of total protein content) from a plasmidic pHT304-18ΩP$_{spo0A}$'*lacZ* fusion was measured in shaken LB at 30°C from t_0 (the end of exponential phase) to t_5. Medians and interquartile ranges of three (t_5) to seven ($t_0 – t_4$) independent experiments are plotted.

Figure 3: The Δ*spo0A* strain takes longer to kill larvae of *Galleria mellonella*.

The lethal time of *Bt* 407 WT and Δspo0A strains was defined as the time span between injection of the inoculum and the time the larvae stopped moving definitively. Tukey boxplots of the lethal time of 63 (WT) and 57 (Δ*spo0A*) larvae are presented.

Figure 4: Spo0A is required for necrotrophism.

CFUs of total bacterial populations were numerated daily after infection of *G. mellonella* by *Bt* 407 WT and Δ*spo0A* at day 0. Measures of three independent experiments and their medians are plotted.

Figure 5: Functional annotation of the *Bt* 407 genome with EggNOG.

The pie chart representing the number of genes per EggNOG category was generated by our in-house script EggnogGenome from BLASTp results against the EggNOG database (see experimental procedures).

Figure 6: Overlap of the *in silico*, *in vitro* and *in vivo* genes datasets.

A venn diagram was built to show logical relations of three sets of genes: genes found *in silico* to harbor at least one putative Spo0A box in their 5' upstream region, genes found to be differentially expressed between *Bt* 407 WT and Δ*spo0A in vitro* (union of the LB t_2 and LB t_5 sets), and genes found to be differentially expressed between *Bt* 407 WT and Δ*spo0A in vivo*. The 184 genes at the intersection of all three sets were defined as the core regulon.

Figure 7: Functional annotation of the Spo0A core regulon.

Genes of the core regulon were clustered according to their functional category(ies). Bars going up represent the number of genes up-regulated in the *Bt* 407 Δ*spo0A* mutant compared to the WT for each category, and bars going down the number of down-regulated genes. When a gene was up-regulated *in vivo* and down-regulated *in vitro* (or the opposite), the *in vivo* condition was arbitrarily used.

Figure 8: Distribution of the core regulon genes according to their Spo0A and *in vivo* regulation.

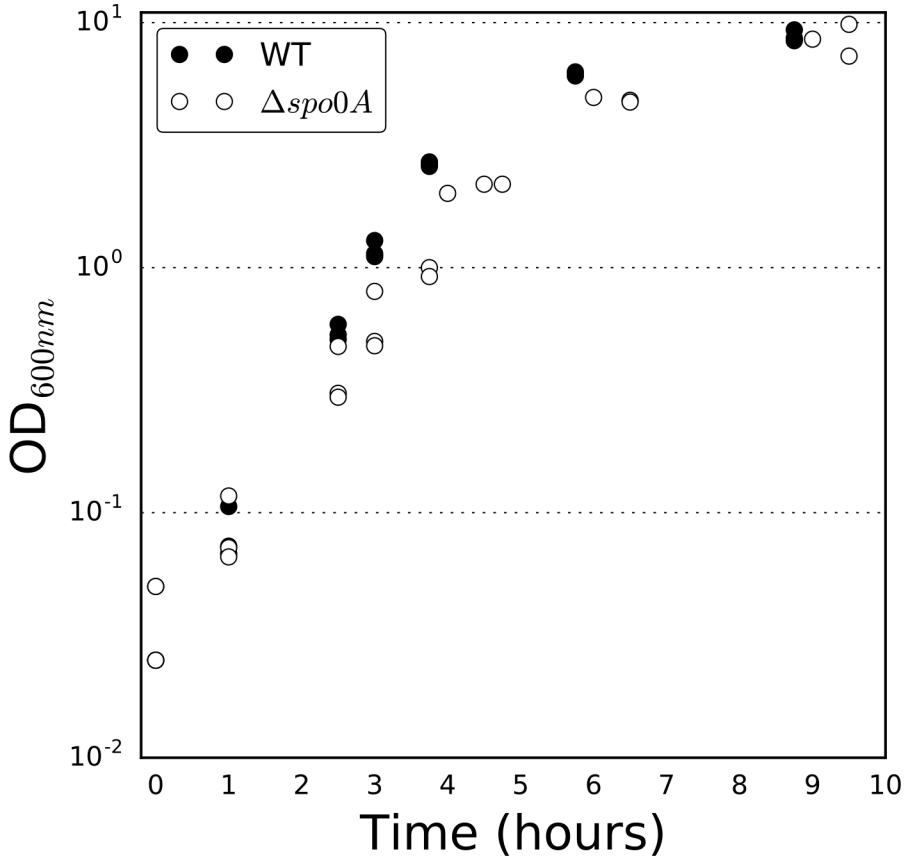
Genes of the core regulon were plotted according to two parameters: their *x* coordinate is their expression ratio between *Bt* 407 Δ*spo0A* and WT strains *in vivo*, and their *y* coordinate is their expression ratio in *Bt* 407 WT between *in vivo* and *in vitro* (LB t_5) conditions.

562 **Figures**

563 Figure 1

564

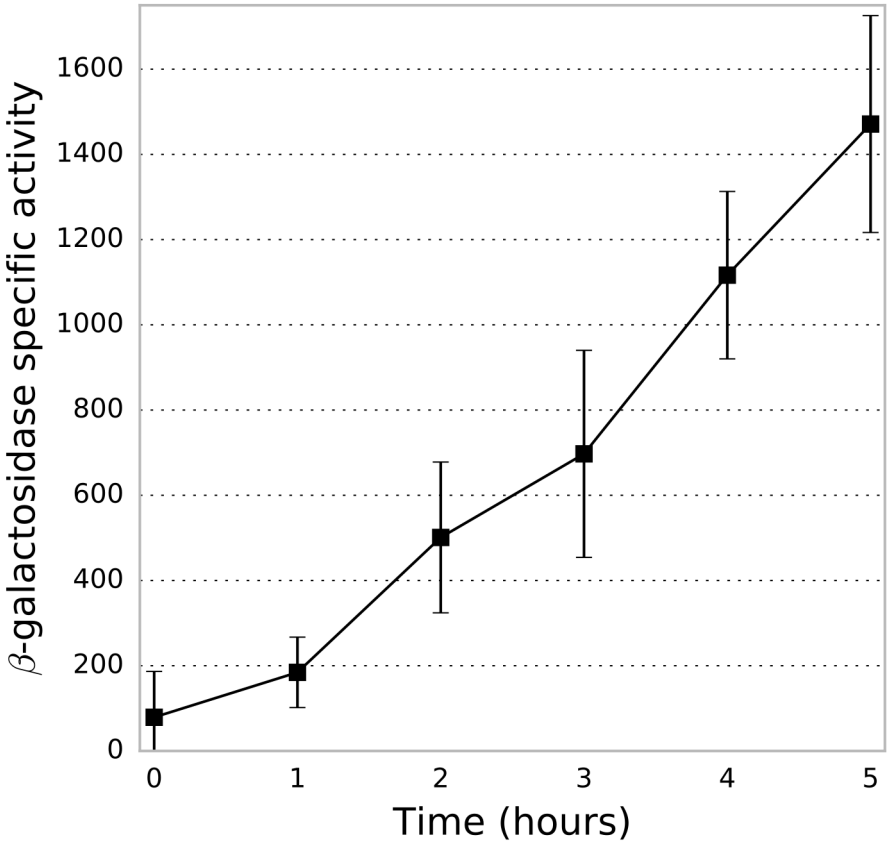
565



566 Figure 2

567

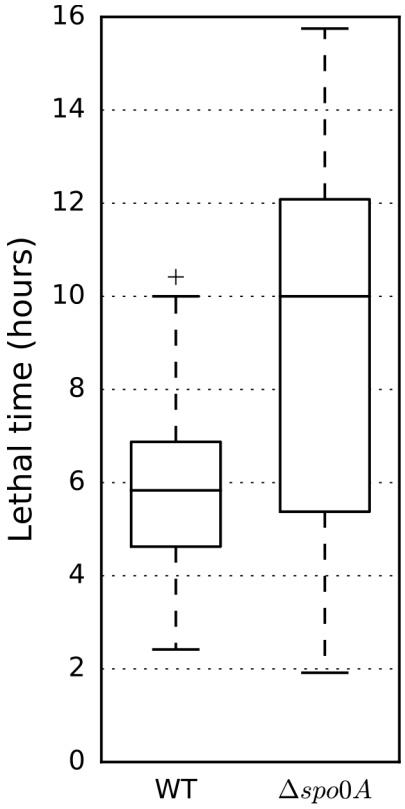
568



569 Figure 3

570

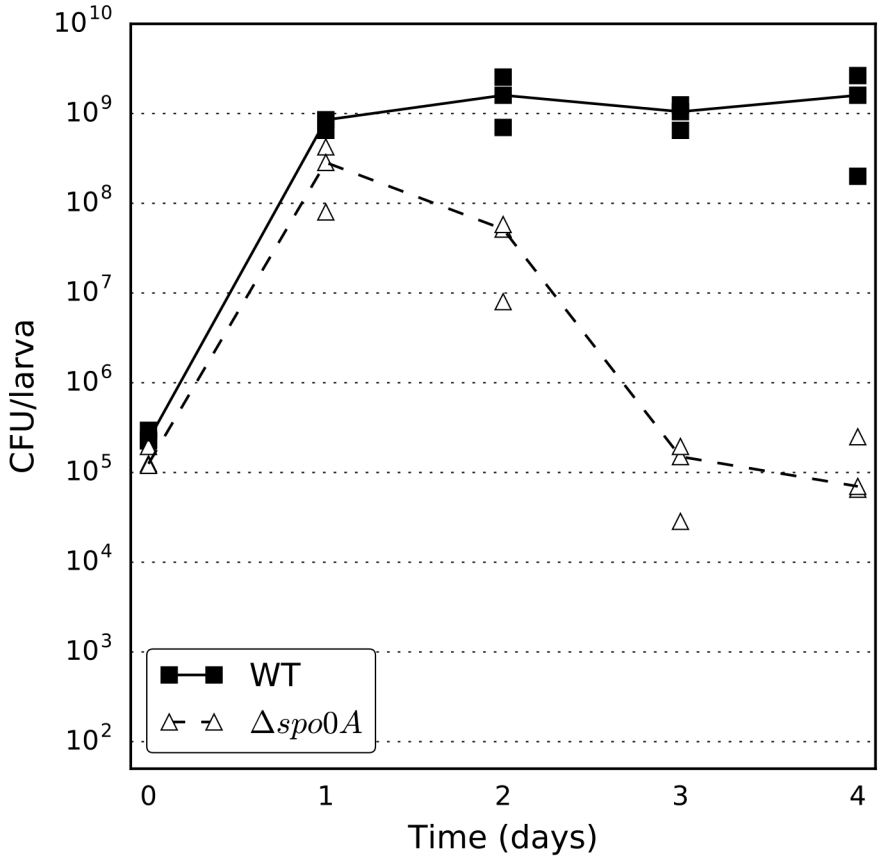
571



572 Figure 4

573

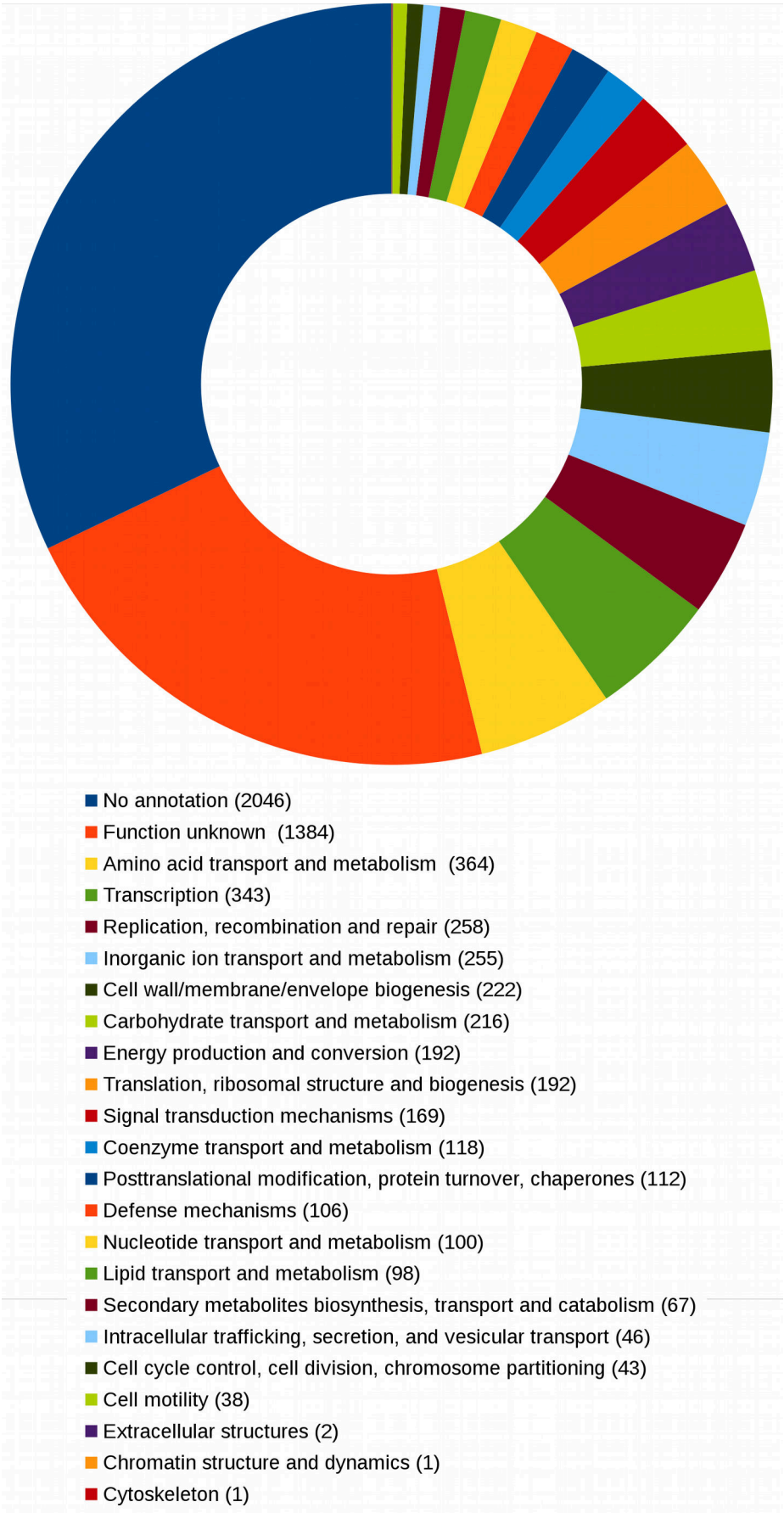
574



575 Figure 5

576

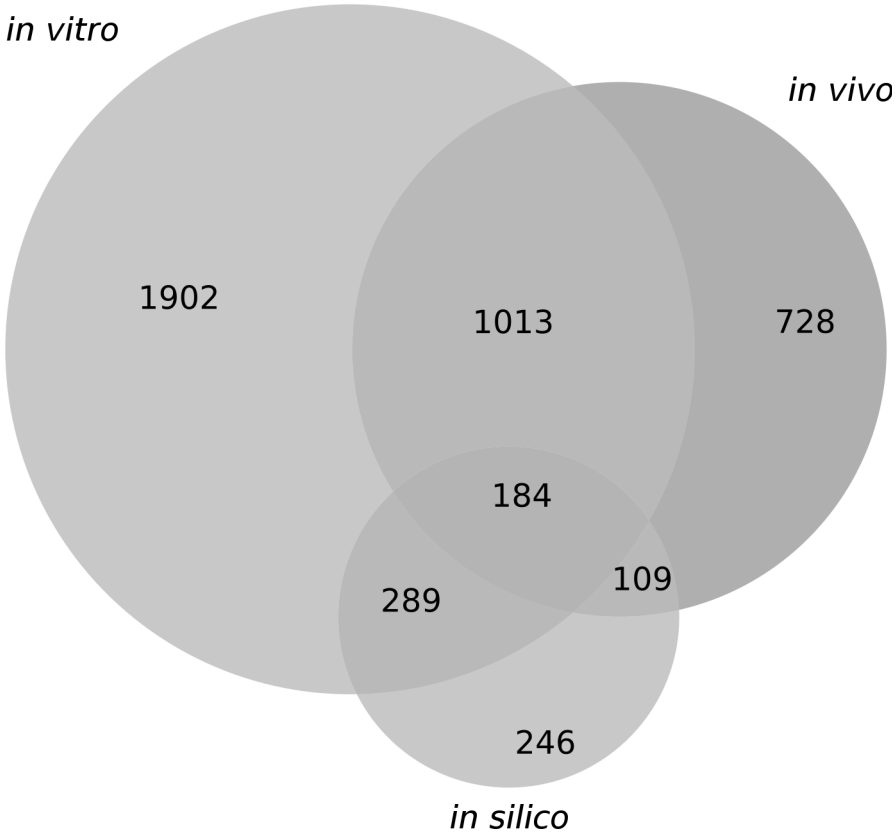
577



578 Figure 6

579

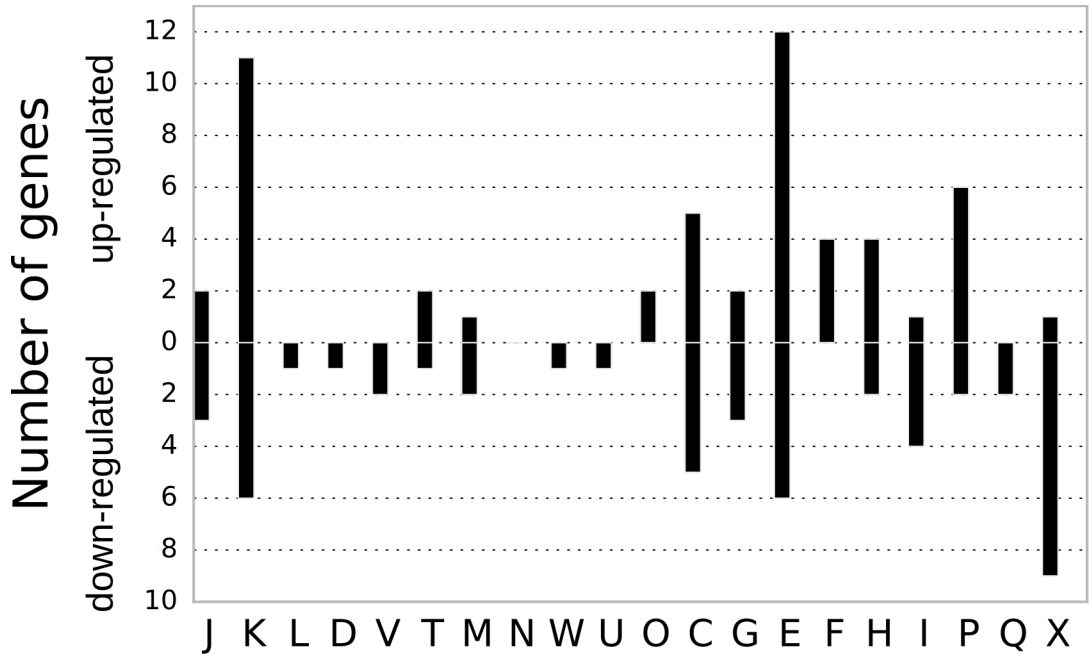
580



581 Figure 7

582

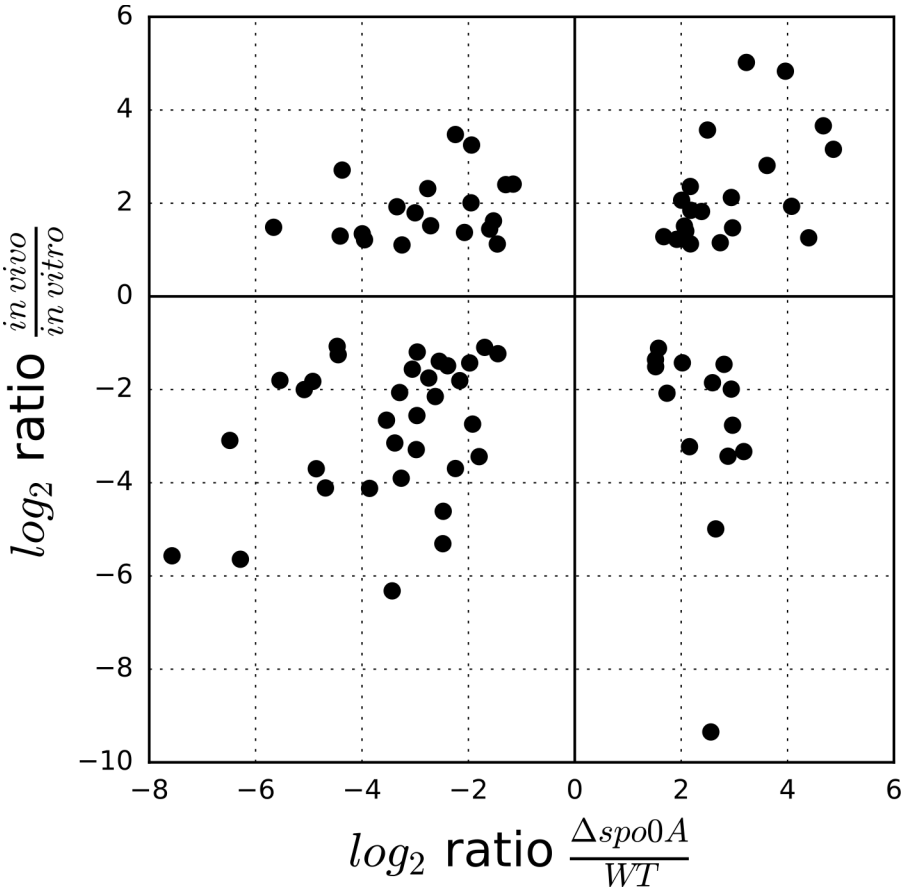
583



584 Figure 8

585

586



Supplementary material

Table S1: strains constructed for this study.

Strain	Reference
Bacillus thuringiensis str. 407 Cry$^-$ (*Bt* 407)	Laboratory collection
Bt 407 *spo0A*::*aphA3*	(Lereclus et al. 1995)
Bt 407/pHT304-18ΩP$_{spo0A}$'*lacZ*	This study

Table S2: plasmids constructed for this study.

Name	Features
pHT304-18Ω*P$_{spo0A}$'lacZ*	The -300 to -14 region upstream of *spo0A*, relative to its predicted ATG, was amplified with SG5 and SG6 on *Bt* 407 genomic DNA. It was then cloned in the pHT304-18Z plasmid between the BamHI and HindIII restriction sites. This construction was used to monitor *spo0A* expression in shaken LB at 30°C.

Table S3: primers used in this study.

Name	Sequence*	Restr. site
SG5	gggaagcttACATCAAGCTGGAATTAATAT	HindIII
SG6	cccggatccGGCTACATGAATTGTTCGAC	BamHI

* Nucleotides not matching the template are lowercase. Restriction sites are underlined.

Table S4: functional re-annotation of the *Bt* 407 genome.

Table S5: the Spo0A core regulon.

Table S6: the Spo0A *in vivo* specific sub-regulon.

603 Table S7: key loci names and their annotation on the *Bt* 407 reference genome

Locus name in this study	Locus name in the reference genome
gerA	BTB_c35820
ynzD	BTB_c13780
spoIIB	BTB_c45930
cytK	*hlgB*
inhA2	*ina1*
nprR	*nprA*
sinI	BTB_c13270
dppE4-oppB3-oppC3-oppD2-oppF3	*opp* operon

604

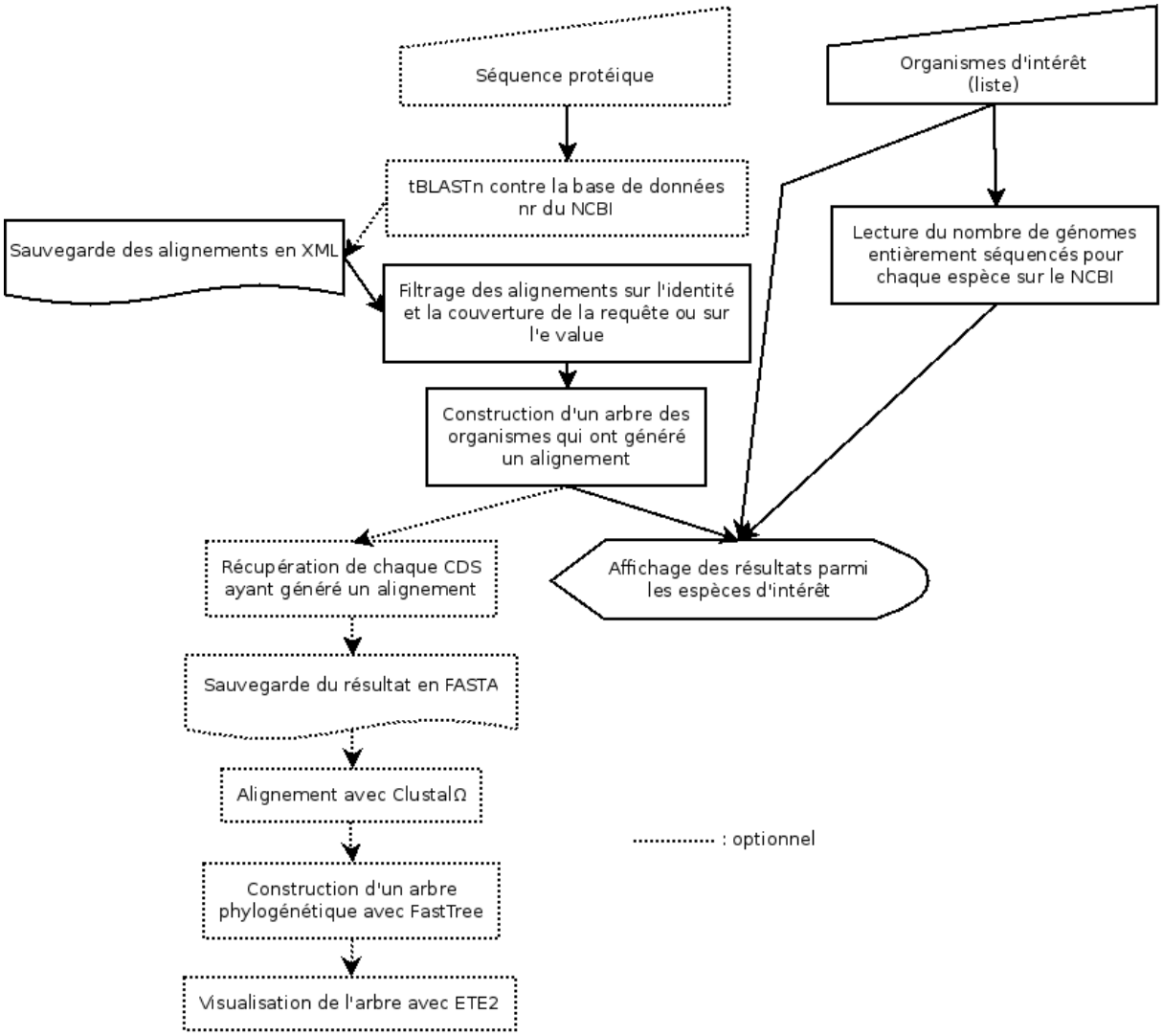


Fig. 25 : Workflow de BLASTats.

Chapitre III - Logiciels et bases de données développés pendant la thèse

1 BLASTats

1.1 Généralités

BLASTats tente de répondre à la question : comment une protéine donnée est distribuée au sein d'un groupe d'espèces d'intérêt ? Autrement formulé, BLASTats donne une idée de la distribution d'homologues d'une protéine dans une liste d'espèces définie. Dans notre cas, la question s'est posée de savoir si *krs*, l'opéron responsable de la synthèse non-ribosomique de la kurstakine, était spécifique du groupe *Bacillus cereus*, et s'il y était réparti de façon homogène. Bien que cette question puisse être abordée de façon « manuelle » en utilisant des interfaces disponibles en ligne, l'automatisation du pipeline a permis de rendre le processus plus fiable et reproductible. Cette automatisation est d'autant plus intéressante que le nombre de génomes disponibles augmente régulièrement, et qu'il peut être judicieux de relancer l'analyse cycliquement.

1.2 Développement

BLASTats a été développé en Python (Python 3 pour 95 %, Python 2 pour 5 %). Le code est organisé en 2 fichiers cumulant 600 lignes. En plus de la librairie standard, il repose sur les modules Biopython (Cock *et al.*, 2009), ETE2 (Huerta-Cepas *et al.*, 2010) et GTK+, et sur les programmes ClustalΩ (Sievers *et al.*, 2011) et FastTree (Price *et al.*, 2009, 2010). Le code est orienté objet et contient 2 classes, une pour l'interface graphique et une contenant les méthodes de traitement des données à proprement parler, une méthode par étape de traitement. Une exécution sur plusieurs fils a été implémentée afin d'accélérer une étape de récupération de données distantes sur le NCBI.

La dernière version du code source de BLASTats est disponible sur GitHub : https://github.com/seb-ksl/BLASTats.

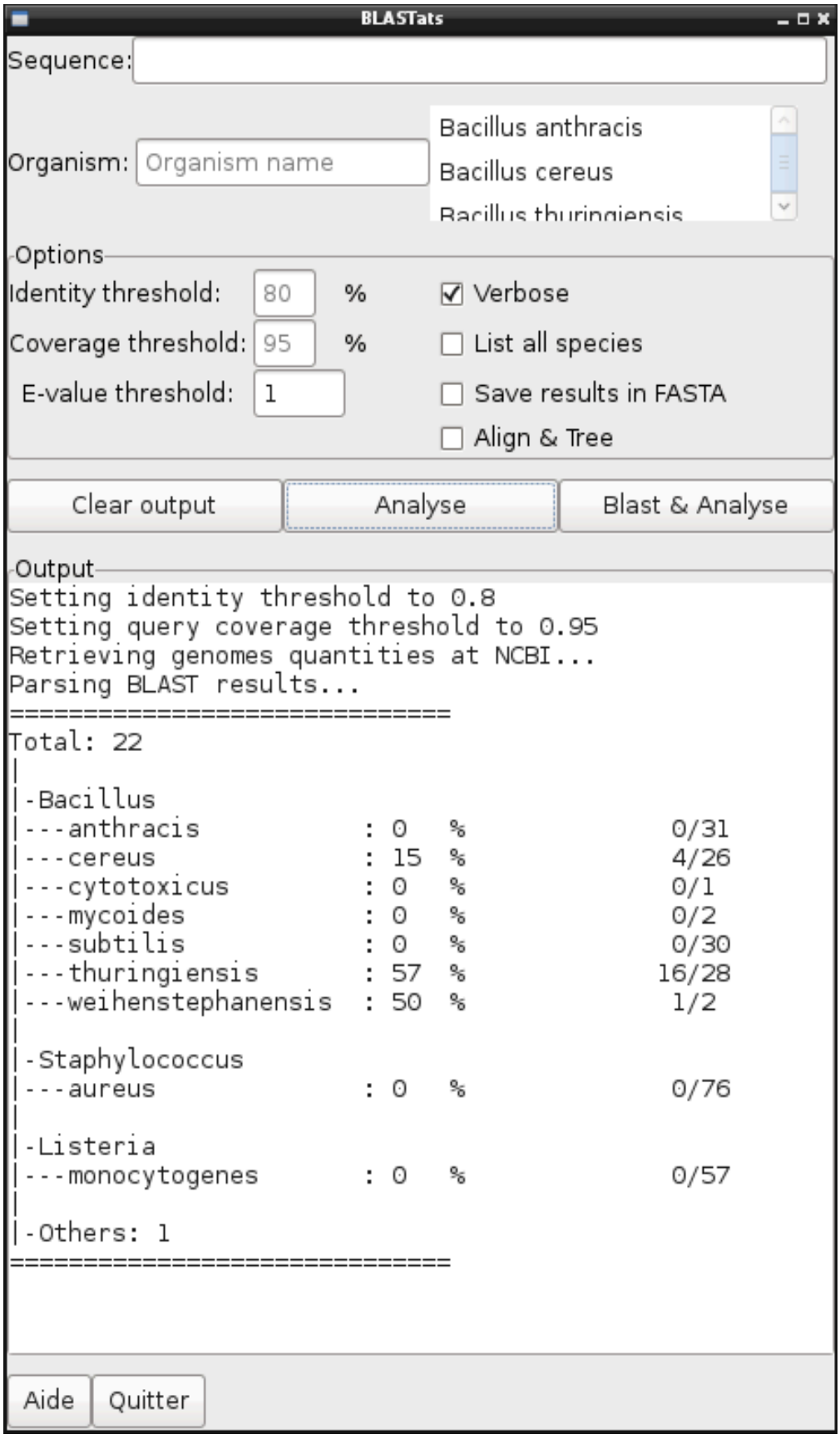


Fig. 26 : Interface de BLASTats.

1.3 Fonctionnement

Un schéma général du workflow de BLASTats est présenté en figure 25. Le programme prend donc en entrée une séquence protéique et une liste d'organismes d'intérêts.

À partir de la liste d'organismes, d'abord convertie en arbre où les genres contiennent des espèces, le programme lit, sur les serveurs du NCBI, le nombre de génomes entièrement séquencés pour chaque espèce. Le choix de ne travailler que sur les génomes complets se justifie dès lors qu'on souhaite inférer des statistiques à partir des alignements. En effet, si on trouve notre protéine d'intérêt dans un génome incomplet, aucun problème ne se pose ; en revanche, si on ne trouve pas la protéine dans un génome incomplet, il est impossible de savoir si la protéine n'y est réellement pas, ou si elle n'a simplement pas été séquencée (ou mal séquencée, ou à cheval sur deux contigs, *etc*.). Les génomes incomplets sont aussi plus susceptibles d'être mal assignés taxonomiquement, particulièrement dans le groupe *Bacillus cereus* où la présence de plasmides est importante. Prendre en compte les génomes incomplets pourrait donc fausser la mesure d'abondance de la protéine dans une espèce. Le revers de ce choix est évidemment la diminution du nombre de génomes pris en compte dans l'analyse. Cette étape est exécutée sur autant de fils parallèles que d'organismes, afin d'optimiser le temps d'attente dans le cas d'une liste d'organismes d'intérêt s'allongeant.

La séquence protéique fournie par l'utilisateur est envoyée aux serveurs du NCBI pour être alignéee par l'algorithme tBLASTn sur les séquences de la base de données nr (pour non-redondantes). Le choix de tBLASTn plutôt que BLASTp permet notamment de détecter des protéines dont le gène n'aurait pas été correctement annoté sur un génome. tBLASTn est appelé avec des paramètres permettant 1) de ne prendre en compte que les génomes complets, 2) d'ignorer les plasmides (afin d'éviter de compter une même souche plusieurs fois), 3) de récupérer 500 alignements (contre 100 par défaut), et 4) de ne conserver que le meilleur HSP (« high scoring segment pair ») de chaque alignement (pour des questions de taille de fichier de résultats). Les résultats du tBLASTn sont téléchargés au format XML et conservés afin de pouvoir refaire l'analyse directement, sans avoir à relancer cette étape.

Le fichier de résultats est lu à l'aide du module Biopython, et pour chaque alignement le meilleur (et seul) HSP récupéré est filtré sur des critères fournis par l'utilisateur. Les critères possibles sont le pourcentage d'identité partagée entre la requête et le sujet (séquence de la base de données s'alignant avec la requête), le pourcentage de couverture de la requête par le sujet, et l'e value calculée par BLAST. Les paramètres par défaut suggèrent des seuils à

80 % d'identité sur 95 % de couverture, des valeurs qui ont montré une bonne fiabilité dans l'annotation de classes enzymatiques par homologie (Rost *et al.*, 2003). À ce stade, le titre des alignements ayant passé le filtre est analysé pour y récupérer le nom de l'espèce et de souche dont le génome a généré l'alignement. Cette liste de souches à une dimension est ensuite reclassée dans un arbre où les genres contiennent des espèces qui contiennent des souches. Cet arbre des alignements est ensuite confronté à l'arbre des espèces d'intérêt et les résultats sont affichés (Fig. 26).

Le programme peut ensuite, optionnellement, continuer en récupérant les séquences codantes complètes ayant généré chaque alignement filtré, et en construisant un alignement global multiple avec ClustalΩ. Cet alignement peut alors être utilisé pour construire un arbre phylogénétique approchant du maximum de vraisemblance à l'aide de FastTree. Cet arbre est finalement visualisé à l'aide du toolkit ETE2.

2 EggnogGenome

2.1 Généralités

EggnogGenome a été développé au cours des analyses RNA-seq présentées dans les résultats de cette thèse. Lors de ces analyses, nous avons eu besoin de classer un grand nombre de gènes en catégories fonctionnelles, et nous avons choisi de suivre les fameuses catégories COGs (pour « cluster of orthologous groups ») (Tatusov *et al.*, 1997). Nous n'avons cependant pas utilisé la base de données COG, plus maintenue et peu pertinente pour les bactéries à Gram positif, mais EggNOG, une base de groupes d'orthologie non-supervisés hiérarchisés dont la dernière version (4) est basée sur les génomes de 3686 espèces (Powell *et al.*, 2014). Une interface web est disponible pour EggNOG, cependant elle n'autorise pas les requêtes multiples et il aurait été beaucoup trop fastidieux de soumettre manuellement les séquences des plus de 6300 protéines prédites du génome de *B. thuringiensis* 407. La base de données a donc été téléchargée et l'analyse faite localement à l'aide d'EggnogGenome.

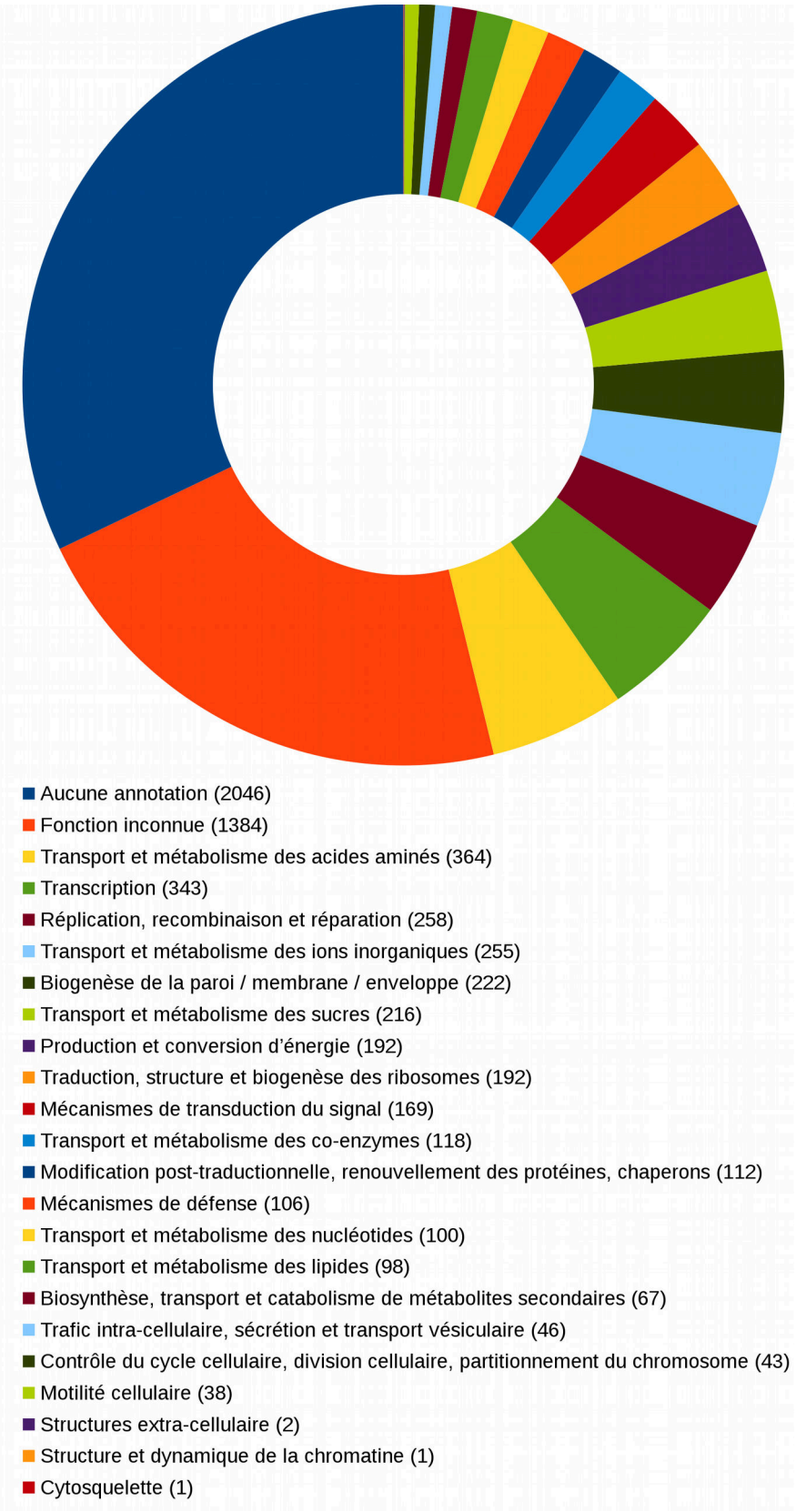


Fig. 27 : Diagramme circulaire des fonctions COG portées par le génome de *Bacillus thuringiensis* 407.

Les données ont été générées par EggnogGenome.

2.2 Développement

EggnogGenome est écrit en 220 lignes de Python 3, et est orienté fonction. Il dépend uniquement, en plus de la librairie standard, de Biopython (Cock *et al.*, 2009). Le code est abondamment commenté (34 % des lignes). Le programme s'utilise dans un terminal, avec des arguments lus par argparse.

La dernière version du code source d'EggnogGenome est disponible sur GitHub : https://github.com/seb-ksl/EggnogGenome.

2.3 Fonctionnement

EggnogGenome vient en bout de pipeline, et présuppose que les étapes suivantes ont été effectuées :

1. Téléchargement du fichier FASTA protéique d'EggNOG et du tableau de correspondance entre protéine et catégorie COG.
2. Transformation du fichier FASTA en base de données BLASTp.
3. BLASTp des protéines du génome à annoter contre la base de données EggNOG créée. Dans notre cas, le programme a été lancé sur un cluster de calcul avec la ligne de commande suivante : `blastall -p blastp -d eggnog.blastdb -i query_proteins.fa -o matches.blast -e .01 -m 7 -a 60 -K 1`. Cette commande permet de fixer un seuil d'e value à 10^{-2} (paramètre -e), d'enregistrer les résultats au format XML (-m 7), d'exécuter le processus sur 60 fils (-a 60) et de ne garder qu'un HSP par alignement (-K 1).

EggnogGenome prend ensuite comme paramètres d'entrée les alignements BLASTp générés et le tableau de correspondance entre protéine et catégorie COG pour assigner à chaque protéine du génome sa ou ses catégorie(s). Les résultats sont enregistrés au format texte à valeurs séparées, et le séparateur peut être choisi entre le point-virgule et la tabulation. Un petit historique contenant le nombre total de protéines alignées, le nombre total de protéines assignées et le nombre de protéines assignées aux catégories R et S (hypothétiques) peut également être généré. Finalement, une quantification des effectifs par catégorie COG peut être enregistré, et permet de présenter l'importance relative des différentes fonctions portées par le génome (Fig. 27).

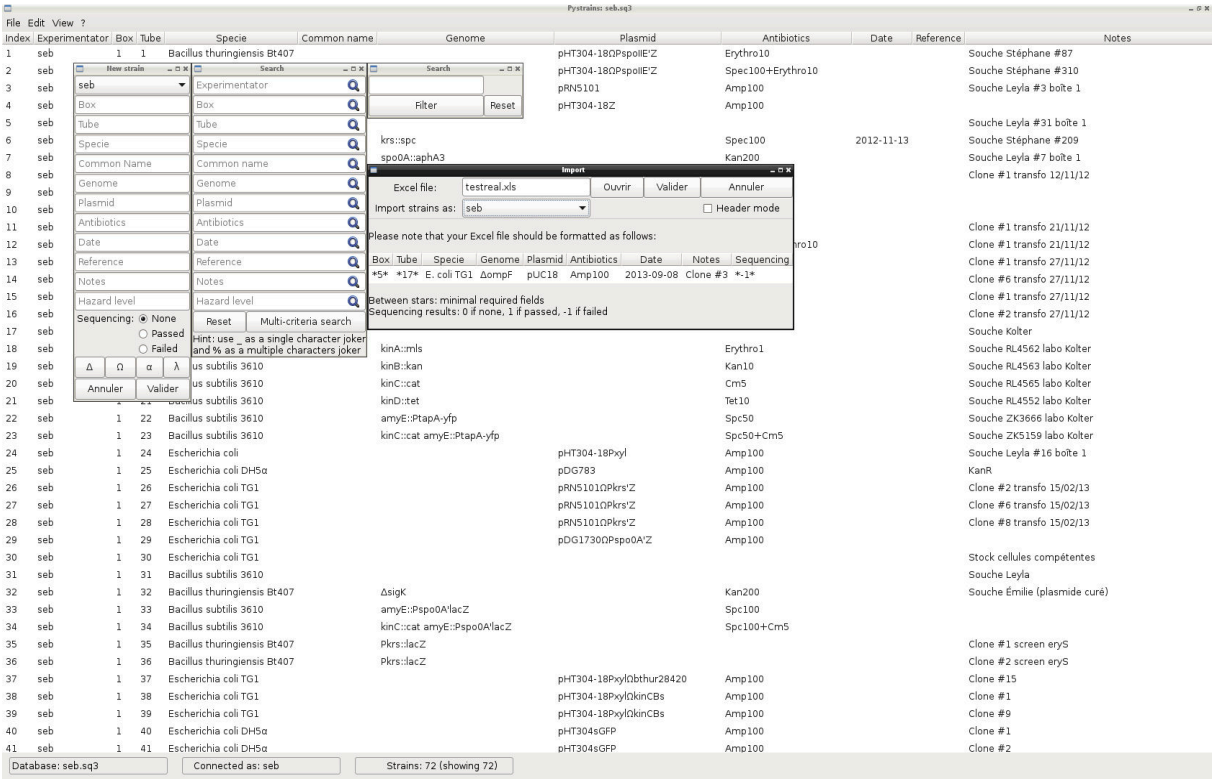


Fig. 28 : Interface de PyStrains.

En fond, la fenêtre principale du souchier. Devant, de gauche à droite, la fenêtre de création d'une nouvelle souche, la fenêtre de filtre avancé, la fenêtre de filtre rapide, et la fenêtre d'import de souches depuis un fichier Excel.

3 PyStrains

3.1 Généralités

Le développement de PyStrains a été démarré pour fournir à l'équipe un moyen efficace de répertorier les souches construites au laboratoire. Il permet une harmonisation des formats de souchiers et une gestion fine des droits d'édition.

3.2 Développement

Le code source de PyStrains représente, à ce stade du développement, plus de 2200 lignes de Python 3. Il est orienté objet, avec une classe pour la gestion de base de données et des classes pour chaque fenêtre de l'interface graphique. PyStrains repose, en plus de la librairie standard, sur les modules GTK+, xlrd et sqlite3. Le choix d'SQLite comme moteur de base de données a été guidé par des considérations techniques, notamment l'absence de personnel et de temps pour maintenir un réel serveur de base de données. SQLite permet donc d'utiliser des fichiers, facilement partageables sur l'infrastructure existante au laboratoire, et lus par PyStrains.

La dernière version du code source de PyStrains est disponible sur GitHub : https://github.com/seb-ksl/PyStrains.

3.3 Fonctionnement

PyStrains fonctionne comme une application classique d'interrogation et édition de base de données. L'interface graphique permet, au premier lancement, de créer le compte administrateur ainsi qu'une première base de données vide. Sans plus de paramétrage, PyStrains peut ainsi être utilisé localement sur un poste de travail, pour gérer son propre souchier (Fig. 28). S'il est prévu un usage commun du souchier, l'administrateur peut alors créer autant d'utilisateurs que nécessaire, et définir leurs droits. Un utilisateur standard n'aura de droits que sur ses propres souches, un encadrant pourra avoir les droits sur ses souches et celles de ses étudiants, et l'administrateur a tous les droits. Les mots de passe des utilisateurs sont chiffrés par SHA-256 puis stockés dans la base de données. La base de données est sauvegardée à chaque clôture de PyStrains, et il est possible de restaurer une sauvegarde en cas de problème.

Pour remplir le souchier, deux possibilités : saisir les souches directement dans

PyStrains, ou importer un souchier existant au format XLS. L'import depuis un fichier XLS nécessite simplement un formatage défini correspondant aux colonnes utilisées par PyStrains. Ces colonnes sont : le nom de l'expérimentateur, le numéro de boîte de stockage, le numéro de tube, le nom de l'espèce, son nom commun, son génotype chromosomique, son génotype plasmidique, les antibiotiques de sélection, la date de construction, une référence bibliographique, des notes, l'état de séquençage et le niveau de danger biologique. Les seuls champs obligatoires pour l'import sont la mention de l'expérimentateur, et pour chaque souche, un numéro de boîte et de tube ainsi que son statut de séquençage.

L'interrogation de la base peut se faire de deux façons : filtre rapide ou filtre avancé. Le filtre rapide recherche un terme dans n'importe quel champ, alors que le filtre avancé permet de rechercher un terme dans un champ donné et de croiser les recherches avec plusieurs termes sur plusieurs champs. Les deux types de recherche permettent l'utilisation de jokers.

4 Bases de données

Au cours du travail de thèse, nous avons rencontré plusieurs situations ayant nécessité la création de petites bases de données très spécifiques.

Un des problèmes fréquemment rencontrés au laboratoire est dû à l'utilisation de trois génomes modèles : le génome de *B. cereus* ATCC 14579 et les deux séquençages du génome de *B. thuringiensis* 407. Les numéros de loci et les noms de gènes n'étant pas harmonisés, il est souvent nécessaire de convertir d'une nomenclature vers l'autre. Le problème s'est accentué lors de la récente ré-annotation automatique des génomes par le NCBI, générant encore une nouvelle nomenclature. Nous avons donc construit plusieurs tableaux de correspondance par BLASTp unidirectionnel, permettant de « traduire » de façon la plus fiable possible un numéro de locus d'une nomenclature vers une autre. Les différents tableaux sont accessibles publiquement à ces adresses : http://gelis.ch/lab/bc2bt/, http://gelis.ch/lab/bt2bc/ et http://gelis.ch/lab/bt2bt/.

Deux autres bases de données ont été créées lors de l'exploitation des données de RNA-seq présentées dans les résultats de cette thèse. De la même façon qu'EggnogGenome a été développé pour assigner une catégorie COG à toutes les protéines du génome de *B. thuringiensis* 407, nous avons souhaité avoir un accès rapide et automatisé aux prédictions

proposées par les programmes HHSearch et InterProScan pour la totalité des séquences codantes du génome de *B. thuringiensis* 407 (Jones *et al.*, 2014; Söding, 2005). Le programme HHSearch a été lancé en boucle sur un cluster de calcul avec la commande `hhsearch -i query.fa -d pdb70/pdb70.hhma -o output/gene.txt -cpu 4 -v 1`. Cette commande permet d'exécuter hhsearch une requête à la fois (stockée dans le fichier query.fa), contre la base de données PDB70 (contenant les modèles de Markov cachés des protéines dont la structure a été résolue et déposée dans la PDB, avec un regroupement des protéines dont les séquences primaires sont identiques à plus de 70 %), sur quatre fils parallèles (-cpu 4) avec une sortie peu verbeuse n'affichant que les avertissements (-v 1). Le programme InterProScan a été également lancé en boucle, mais cette fois directement sur les serveurs publics. Pour cela, l'EBI fournit des scripts en différents langages permettant d'utiliser le service *via* une API SOAP ou REST. Dans notre cas c'est l'API SOAP qui a été sollicitée en boucle par la commande `iprscan5_soappy.py --goterms --nopathways --email=sgelis@jouy.inra.fr --title=foo --outfile=bt407_data/gene query.fa`. Le paramètre --goterms permet d'inclure les données GO (pour « gene ontology ») telles que le processus biologique, la fonction moléculaire et le compartiment cellulaire, et le paramètre --nopathways permet de ne pas inclure les détails sur les voies métaboliques. Les bases de données résultantes sont accessibles publiquement à ces adresses : http://gelis.ch/lab/iprscan/ et http://gelis.ch/lab/hhsearch/. Les informations pour chaque gène sont accessibles directement par URL simple, par exemple http://gelis.ch/lab/iprscan/lgrC/ ou http://gelis.ch/lab/hhsearch/lgrC.txt, ou à travers le formulaire de la page d'accueil en renseignant le nom du gène ou le numéro de locus.

Discussion et perspectives

Chapitre I - La kurstakine

1 La régulation de la production de kurstakine

1.1 Régulation transcriptionnelle de *krs*

La transcription de l'opéron *krs* se révèle être complexe, faisant intervenir au moins deux régulateurs majeurs de *B. thuringiensis* (NprR et Spo0A) ainsi qu'une séquence palindromique. Il semble également que les conditions de culture aient un effet important sur la force et le timing d'expression du promoteur. Ces effets sont maintenant bien décrits mais encore mal expliqués, et de nombreux travaux permettraient une meilleure compréhension des mécanismes de régulation de la transcription de *krs*.

1.1.a Le rôle exact du palindrome

Pour mieux comprendre la transcription de *krs*, il faudrait en premier lieu élucider le rôle exact du palindrome que contient son promoteur. La séquence composant ce palindrome, deux répétitions exactes tête-bêche de 15 nt séparées par 6 nt, a un effet important sur la transcription de *krs*, mais nous n'avons encore aucune hypothèse quant au mécanisme impliqué. Cette région n'étant apparemment pas transcrite, au moins dans nos conditions de travail, il semble exclu qu'il s'agisse d'un riboswitch (Mellin and Cossart, 2015), d'un ARN anti-sens ou même d'un terminateur.

Il faudrait dans un premier temps déterminer si cette séquence agit seule, comme *cis*-régulateur, ou est partenaire d'un *trans*-régulateur. Pour rechercher un partenaire du palindrome on pourrait tenter, par des expériences de biochimie, de co-purifier le palindrome avec son partenaire hypothétique. Étant donné que NprR-NprX et la répétition 5' du palindrome sont requises pour la transcription de *krs*, on peut supposer que le palindrome est en fait le site de fixation de NprR-NprX, et cette hypothèse peut être parmi les premières testées. En revanche, elle n'explique pas comment NprR-NprX régulerait les autres gènes de son régulon, dont aucun promoteur ne semble contenir le palindrome.

1.1.b Les *trans*-régulateurs

Deux régulateurs importants interviennent dans la transcription de *krs* : NprR et

Spo0A. L'activation de la transcription de *krs* requiert la présence de NprR et de son peptide de signalisation NprX (Dubois *et al.*, 2012). Spo0A réprime la transcription de *krs*, probablement au moins partiellement en se fixant à la boîte Spo0A présente entre le démarrage de transcription et le codon d'initiation de la traduction du premier gène de l'opéron. L'intervention de ces deux régulateurs pourrait suffire à expliquer le profil de transcription de *krs* dans le temps en HCT à 30°C : démarrage de transcription après t_1 dépendant de l'activation par NprR (la transcription de *nprR* démarrant à t_1 à 37°C (Dubois *et al.*, 2013)), maximum à t_3 puis baisse progressive résultant de l'accumulation de Spo0A actif dans la cellule. Cependant ce modèle simple présente une contradiction : l'expression de *krs* est plus forte en HCT qu'en LB alors que Spo0A actif, répresseur de P_{krs}, s'y accumule plus rapidement. Pour réconcilier ces résultats, on peut imaginer qu'en HCT, un activateur transcriptionnel de *krs* encore inconnu serait plus actif qu'en LB. Cet activateur pourrait même être sous le contrôle de Spo0A, auquel cas Spo0A activerait alors l'expression d'un activateur transcriptionnel de *krs* tout en réprimant directement son expression *via* la boîte Spo0A décrite.

Il semble donc nécessaire de rechercher plus largement les régulateurs transcriptionnels ayant un effet sur la transcription de *krs*. Une autre étude suggère par exemple que, dans d'autres conditions, SinR pourrait avoir un rôle de répresseur sur la transcription de *krs*, rôle que nous n'avons jamais pu confirmer (Fagerlund *et al.*, 2014). Des données préliminaires acquises dans le cadre de cette thèse montrent par ailleurs qu'au moins deux autres régulateurs impliqués dans la formation de biofilm et la motilité (M. Gohar, données non-publiées) auraient un rôle dans la régulation de la transcription de *krs*. La recherche globale des régulateurs affectant l'expression de *krs* pourrait permettre de mieux expliquer son profil de transcription et pourrait mettre en évidence la ou les connexion(s) entre les conditions environnementales et l'expression de *krs*. Il faudrait également déterminer pour tous ces régulateurs, y compris NprR et Spo0A, s'ils se fixent directement sur le promoteur de *krs* ou si la régulation est indirecte.

1.2 Régulation post-transcriptionnelle

En plus de cette régulation transcriptionnelle riche, et étant donné la complexité de la synthèse de peptides non-ribosomique, on peut supposer que la production de kurstakine est soumise à de nombreuses régulations post-transcriptionnelles. La première, la plus évidente,

est la stabilité de l'ARNm *krs*. L'organisation des gènes en opéron montrée par RT-PCR sur les régions intergéniques et par les données RNA-seq suggère l'existence d'une transcription de tout le locus démarrant dès l'amont du premier gène. Il faut donc imaginer, soit que l'ARNm complet existe effectivement, ce qui nécessiterait une stabilité importante pour éviter la dégradation, soit que la traduction et la dégradation de l'ARNm sont concomitantes. Dans les deux cas, cela représente une régulation importante précédant la production de kurstakine qui n'a pas encore été abordée.

L'assemblage et le fonctionnement du complexe synthétase est un autre élément de régulation commun à tous les NRPS. Dans notre travail sur la kurstakine, ces points n'ont pas été abordés, et il serait intéressant d'acquérir des données enzymatiques sur l'assemblage de ce complexe, sa demi-vie, sa vitesse de réaction, la disponibilité relative des substrats, *etc*.

La sécrétion de la kurstakine est une étape importante consécutive à sa synthèse. Les expériences effectuées par nos collaborateurs de l'équipe ProBioGEM (voir le paragraphe ultérieur sur les applications de la kurstakine) ont très tôt souligné l'impossibilité de détecter la kurstakine dans le milieu de culture de *B. thuringiensis*. Elle n'était retrouvée qu'en très faibles quantités à la surface des spores, laissant supposer qu'elle était soit sécrétée puis rapidement dégradée, soit qu'elle était sécrétée puis associée à la paroi, soit qu'elle n'était pas sécrétée du tout. Des résultats récents, acquis en collaboration avec ProBioGEM, confirment finalement sa sécrétion *via* KrsE. En effet, la surexpression de *krsE* depuis un plasmide a permis d'augmenter la concentration de kurstakine dans le milieu de culture, suffisamment pour la rendre détectable. Inversement, la délétion de *krsE* affecte fortement la viabilité des colonies sur gélose, qui lysent en 48 h, suggérant que l'accumulation de kurstakine intra-cellulaire serait toxique pour la cellule.

Puisque la kurstakine est sécrétée dans le milieu extra-cellulaire, il serait intéressant d'y étudier son devenir. Il est en effet possible, étant donné le nombre important d'enzymes dégradatives sécrétées par *B. thuringiensis* dans son milieu, que la kurstakine y soit rapidement dégradée. Il est donc nécessaire de mesurer sa stabilité dans différents tampons ainsi que dans le milieu de culture conditionné de *B. thuringiensis*. Nous avons par ailleurs construit des mutants de délétion de deux enzymes ayant le potentiel de dégrader la kurstakine, une lactonase et une acyl-hydrolase, mais ces mutants n'ont pas permis d'augmenter la concentration de kurstakine détectée dans le milieu de culture.

2 Le mode d'action de la kurstakine dans la structuration du biofilm

Nous avons pu observer que, comme d'autres lipopeptides, la kurstakine est impliquée dans la structuration du biofilm de *B. thuringiensis* 407 (Raaijmakers *et al.*, 2010). Le rôle exact des lipopeptides dans la formation de biofilms n'est pas encore élucidé : certains promeuvent la formation du biofilm, d'autres la réduisent, d'autres ont un effet sur l'architecture. Dans notre cas, nous avons montré qu'une souche déficiente pour la production de kurstakine forme un biofilm de densité réduite en tube de verre, et un biofilm à la structure fortement affectée en plaque PVC. Le biofilm formé en tube de verre est probablement lui aussi affecté dans sa structure, mais la mise en évidence expérimentale est difficile dans ce système. La différence de structure observée se traduit aussi en termes mécaniques. En effet, un biofilm formé par la souche parentale *B. thuringiensis* 407 est percé par un instrument contondant (en l'occurrence une pipette Pasteur fondue pour obtenir une bille de verre à son extrémité) alors que le biofilm formé par un mutant ne produisant pas de kurstakine est emporté intact au fond du puits lorsque la pression est appliquée.

Au moins deux hypothèses non-mutuellement exclusives permettent de faire le lien entre la kurstakine et le biofilm. La première serait liée directement aux caractéristiques physico-chimiques supposées de la molécule. Sa nature amphiphile pourrait en effet lui permettre de jouer un rôle sur la tension de surface et pourrait faciliter l'adhésion aux surfaces, donc la formation de biofilm. La seconde hypothèse placerait la kurstakine comme molécule de signalisation, à l'instar de la surfactine (López *et al.*, 2009b). La surfactine, pour rappel, active la phosphorylation de Spo0A *via* la kinase KinC chez *B. subtilis*, ce qui a notamment pour effet de permettre la formation de biofilm par le biais de l'antagonisation de SinR par SinI. Nous avons étudié cette hypothèse, et le premier obstacle a été l'absence d'analogue de KinC chez *B. thuringiensis* 407. Faute de pouvoir trouver un analogue en se basant sur la séquence primaire ou la structure secondaire de KinC, nous avons tenté de mettre en place une approche génétique consistant à cribler une banque génomique d'expression de *B. thuringiensis* 407 dans une souche rapporteur de *B. subtilis*. Dans cette souche de *B. subtilis*, les gènes codant pour les kinases KinB et KinC sont interrompus et le promoteur soit de *tapA*, soit de *spo0A* est fusionné à *lacZ*. Si cette souche est cultivée en présence de surfactine, elle doit être incapable d'en sentir le signal. Un clone transformé par un vecteur de la banque de *B. thuringiensis* 407 contenant un gène codant pour un analogue fonctionnel de KinC devrait en revanche exprimer le gène rapporteur. Malheureusement cette

expérience n'a jamais pu être mise au point, à cause de problèmes de fuite d'expression des fusions *lacZ* : toutes les colonies étant bleues, impossible d'utiliser ces souches pour cribler une banque.

La différence mécanique observée entre les biofilms formés par la souche *B. thuringiensis* 407 parentale et par la souche mutante ne produisant pas de kurstakine peut également être expliquée de plusieurs façons. La pellicule formée par la souche mutante pourrait être plus élastique, plus résistante à la force de pression perpendiculaire. Si cette force de cohésion est inférieure à la force d'adhésion au PVC dans la souche parentale et devient supérieure dans la souche mutante, cela pourrait expliquer nos observations : la pellicule formée par le mutant se détache du support sans se déchirer alors que celle formée par la souche parentale se déchire avant sans se détacher. Le résultat serait identique si la force d'adhésion au support de la souche mutante était diminuée par rapport à la souche parentale, et il est impossible pour le moment de trancher entre ces deux hypothèses, qui ne sont de toute façon pas mutuellement exclusives. Il faudrait pour cela se tourner vers des outils de rhéologie qui permettraient d'évaluer la résistance et l'élasticité des deux pellicules. Il faudrait également tester un éventuel défaut d'adhésion aux supports de la souche mutante.

3 Les autres rôles de la kurstakine

La nature chimique de la kurstakine et les travaux menés sur d'autres lipopeptides laissent supposer qu'elle pourrait jouer de multiples rôles au cours du cycle de vie de *B. thuringiensis* (Raaijmakers *et al.*, 2010). Nous avons, de plus, pu montrer que l'opéron responsable de la synthèse de kurstakine est statistiquement plus présent dans les génomes de *Bt* (45 %, 9/20) que dans ceux de *Bc* (12 %, 3/26), ce qui suggère une possible implication de la kurstakine dans le mode de vie spécifique de *B. thuringiensis*, c'est-à-dire son entomopathogénicité. Des données récentes abondent en ce sens et montrent que si la kurstakine n'est pas nécessaire au nécrotrophisme de *B. thuringiensis* dans *Galleria mellonella*, la sur-expression de *krs* restaure quand même partiellement le nécrotrophisme dans une souche mutante de délétion de *nprR-nprX*, normalement incapable de survivre dans le cadavre de l'hôte (Dubois *et al.*, 2012).

On peut notamment envisager que la kurstakine puisse servir de molécule anti-microbienne, permettant à *B. thuringiensis* d'entrer en compétition active avec la flore

commensale de son hôte insecte. Des activités anti-bactériennes, anti-fongiques et anti-virales ont déjà été associées à d'autres lipopeptides et pourraient être testées avec la kurstakine (Cameotra and Makkar, 2004; Schneider *et al.*, 2014). Des données préliminaires datant de la découverte de la kurstakine suggèrent qu'elle pourrait effectivement avoir une action anti-fongique (Hathout *et al.*, 2000).

Une autre possibilité, encore inédite pour un lipopeptide, serait que la kurstakine permette à *B. thuringiensis* de se défendre contre l'immunité cellulaire de l'hôte. Cette hypothèse est rendue plausible par des résultats montrant qu'un lipopeptide de *Pseudomonas fluorescens* est actif contre une amibe, et que l'expression du locus NRPS responsable de la synthèse de ce lipopeptide est induite par la présence de l'amibe (Mazzola *et al.*, 2009). Les amibes et les cellules immunitaires phagocytaires partageant un même mode d'action, on peut imaginer que certains lipopeptides puissent être actifs contre ces derniers, qui représentent une part importante de l'immunité des invertébrés (Little *et al.*, 2005). Cependant, tout reste à faire pour tester cette hypothèse, à commencer par la reproduction des résultats de Mazzola *et al.* (2009) avec la kurstakine. On peut ensuite imaginer étendre les résultats obtenus sur amibes à des cellules phagocytaires.

Enfin, la kurstakine pourrait agir simplement comme agent tensioactif permettant à *B. thuringiensis* de se propager plus efficacement sur milieu solide. La surfactine a en effet été montrée comme impliquée dans la mobilité par sliding et par swarming (Grau *et al.*, 2015; Julkowska *et al.*, 2005). Si la kurstakine permet également d'améliorer la mobilité sur les surfaces, on peut imaginer qu'elle conférerait un avantage lors de l'infection, en permettant une meilleure dissémination dans l'hôte.

Quel que soit le rôle potentiel de la kurstakine au cours de l'infection de *Galleria mellonella*, il serait intéressant de suivre sa transcription dans la larve après inoculation. Même si cette expérience nécessite encore beaucoup de mise au point, on peut envisager de fusionner le promoteur de *krs* à un rapporteur luminescent, ce qui permettrait d'observer la cinétique d'expression de *krs* au cours de l'infection d'une même larve, sans la sacrifier.

3.1 Applications

Une collaboration a été développée avec le laboratoire ProBioGEM (Polytech Lille et Université Lille I), avec pour objectif l'amélioration de la production de kurstakine en vue de

permettre son exploitation industrielle. Cette collaboration a permis l'obtention de résultats innovants, et un brevet est en cours de dépôt. Certains résultats concernant l'amélioration de la sécrétion de kurstakine ont déjà été présentés dans un paragraphe précédent. Les autres résultats concernent d'autres aspects de la construction d'une souche sur-productrice de kurstakine, notamment le choix d'une souche de départ ayant un bon niveau de production, l'augmentation de la transcription de *krs* et l'augmentation de l'activité du complexe synthétase. Les procédés de production et de purification sont toujours en cours d'amélioration.

Chapitre II - Spo0A chez *Bacillus thuringiensis*

1 Les sites réels de fixation de Spo0A sur l'ADN

Dans notre travail sur le régulon Spo0A, nous nous appuyons en partie sur une prédiction *in silico* des boîtes Spo0A dans le génome de *B. thuringiensis* 407. Cette prédiction se base sur des données acquises chez *B. subtilis* et nous faisons donc deux hypothèses : la boîte Spo0A est conservée entre *B. thuringiensis* et *B. subtilis*, et le consensus de la boîte Spo0A est suffisamment robuste pour toutes les détecter dans le génome. La première hypothèse est raisonnable, car la séquence en acides aminés de Spo0A est très conservée entre *B. thuringiensis* et *B. subtilis* : 81 % d'identité stricte et 89 % de similarité sur toute la séquence, 92 % d'identité et 97 % de similarité sur le domaine C-terminal de liaison à l'ADN (Rowe-Magnus and Spiegelman, 1998). De plus, la protéine Spo0A~P purifiée de *B. subtilis* semble capable de se lier à des boîtes Spo0A de *B. thuringiensis in vitro* (Lereclus *et al.*, 2000). En revanche, la seconde hypothèse est plus difficile à évaluer. Chez *B. subtilis* même, certaines boîtes très éloignées du consensus ont été montrées comme fonctionnelles, celles de *spoIIGA* par exemple (Baldus *et al.*, 1994). À l'inverse, une boîte canonique n'est pas garantie d'être fonctionnelle dans la cellule.

Pour confirmer et affiner nos résultats, il serait donc très intéressant de déterminer les sites de fixation réels de Spo0A sur le génome de *B. thuringiensis* 407. Pour cela, des expériences de ChIP-seq seraient envisageables et pourraient remplacer, ou venir en renfort des prédictions *in silico*.

2 Les régulations spécifiques à *B. thuringiensis*

2.1 La sporulation

Nous avons pu mettre en évidence plusieurs différences de régulation de gènes de sporulation entre *B. thuringiensis* et *B. subtilis*, et ces différences pourraient être intéressantes à caractériser. Le promoteur de *spo0F*, par exemple, comporte deux boîtes Spo0A fonctionnelles chez *B. subtilis* (Strauch *et al.*, 1993), alors qu'il n'en contient aucune chez *B. thuringiensis* chez qui il semble néanmoins régulé indirectement par Spo0A. À l'inverse quatre gènes, *ynzD*, *spoIIB*, *spoIID* et *spoIIQ*, sont régulés potentiellement directement par Spo0A chez *B. thuringiensis* mais pas chez *B. subtilis*.

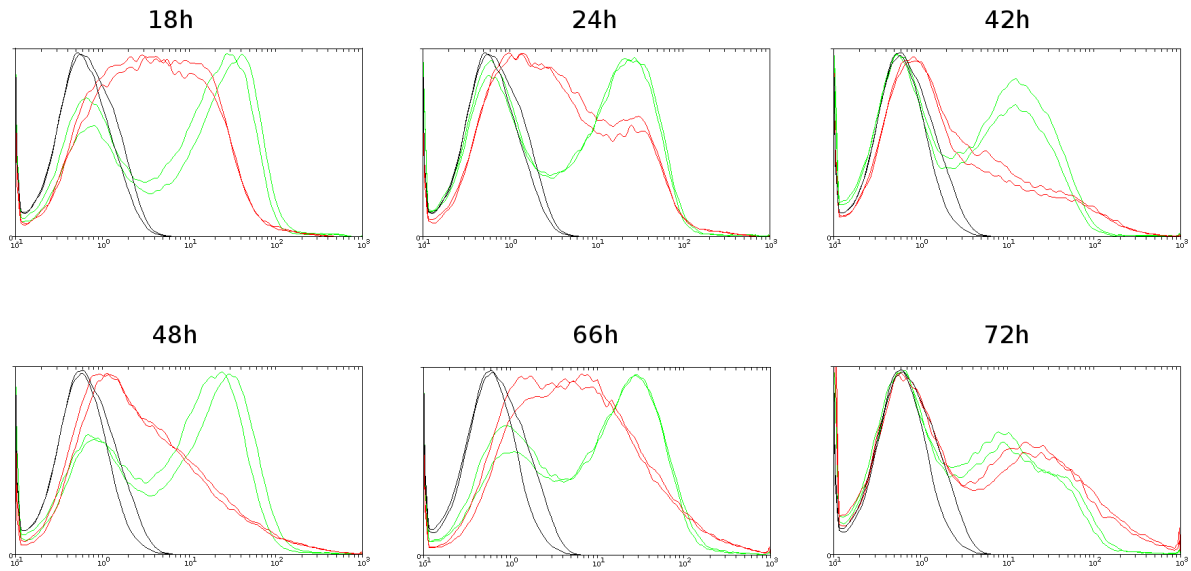


Fig. 29 : Spo0A régule l'expression de *sinI* dans la population.

La fluorescence générée par une fusion plasmidique entre le promoteur de *sinI* et le gène codant pour la protéine fluorescente mCherry a été mesurée au cytomètre en flux. La fluorescence a été suivie dans des biofilms cultivés sur milieu solide LB agar à 30°C pendant 72 h. En noir, deux biofilms indépendants d'une souche parentale sans rapporteur. En vert, deux biofilms indépendants d'une souche parentale avec rapporteur. En rouge, deux biofilms indépendants d'une souche Δ*spo0A* avec rapporteur.

Indépendamment de nos données sur le régulon Spo0A, *B. thuringiensis* possède au moins trois autres spécificités majeures : son gène codant pour le facteur σ^K de sporulation n'est pas régulé par l'élément skin (Stragier *et al.*, 1989), NprR est impliqué dans l'initiation de la sporulation (Perchat *et al.*, article soumis), et les facteurs σ^E et σ^F de sporulation semblent être dupliqués (incluant une copie potentielle de σ^E sur un plasmide).

Tous ces éléments suggèrent que, bien que *B. subtilis* et *B. thuringiensis* soient deux *Bacilli* sporulants relativement proches, de nombreuses différences de régulation existent entre les deux bactéries. Il serait donc intéressant de caractériser davantage la cascade de sporulation de *B. thuringiensis*.

2.2 La formation de biofilm

Spo0A permet la formation de biofilm chez *B. subtilis*, en levant la répression de SinR et AbrB sur les opérons nécessaires à la production de matrice extra-cellulaire (Vlamakis *et al.*, 2013). Le génome de *B. thuringiensis* 407 contient quatre copies du gène *abrB*, et seule l'implication d'*abrB1* dans la formation de biofilm a été déterminée (Fagerlund *et al.*, 2014). Or nous avons pu noter que si seul *abrB1* fait effectivement partie du régulon Spo0A central, *abrB2*, *abrB3* et *abrB4* sont néanmoins régulés indirectement par Spo0A *in vitro*, ainsi qu'*in vivo* pour *abrB3*. Il est d'ailleurs intéressant de constater que la transcription d'*abrB1* et *abrB4* est réprimée par Spo0A alors qu'il active celle d'*abrB2* et *abrB3*. Sachant cela, le rôle des AbrB de *B. thuringiensis* dans la formation de biofilm pourrait donc être abordé.

Une autre différence majeure entre les régulons Spo0A de *B. thuringiensis* et de *B. subtilis* est l'absence de *sinI* dans le régulon Spo0A central de *B. thuringiensis*. L'expression de *sinI* n'a été trouvée activée par Spo0A qu'*in vitro* à t_2 (d'un facteur 3,5), et aucune boîte Spo0A ne semble présente dans son promoteur, suggérant que cette régulation est indirecte. Des données préliminaires d'expression d'une fusion rapporteur entre le promoteur de *sinI* et le gène codant pour la protéine fluorescente mCherry observée au cytomètre en flux semblent indiquer que la régulation transcriptionnelle de *sinI* chez *B. thuringiensis* est complexe. Dans les conditions testées (le biofilm sur milieu solide LB agar incubé à 30°C), la transcription de *sinI* était effectivement diminuée à certains temps dans la souche *B. thuringiensis* 407 Δ*spo0A* comparée à la souche parentale, mais c'est surtout le profil de transcription dans la population qui était affectée : la souche parentale présentait une bimodalité perdue par la souche Δ*spo0A*. La voie de régulation impliquant

Spo0A, SinI et SinR dans la formation de biofilm mériterait donc également d'être davantage étudiée.

2.3 La mobilité

Dans notre étude, nous nous sommes focalisés sur les fonctions et UT potentiellement directement régulées par Spo0A. Or, si l'on s'intéresse également aux fonctions indirectement régulées par Spo0A, il apparaît que la quasi-totalité des gènes de mobilité (92 %, 35/38) est sur-exprimée dans la souche Δ*spo0A* comparée à la souche parentale, dans au moins une condition expérimentale. Cela traduit donc une répression importante de la mobilité par Spo0A, par une voie *a priori* indirecte. Or, parmi les gènes du régulon Spo0A central se trouve *mogR*. MogR est impliqué dans la répression de la mobilité chez *Listeria monocytogenes* (Shen and Higgins, 2006), et sa potentielle implication dans la mobilité de *B. thuringiensis* est suggérée par la conservation de la séquence en acides aminés entre les deux modèles, notamment de ceux impliqués dans la liaison à l'ADN (Shen *et al.*, 2009). La transcription de *mogR* est activée par Spo0A, et il est donc envisageable que Spo0A réprime par son intermédiaire la mobilité de *B. thuringiensis*.

Ces résultats sont en contradiction apparente avec ceux présentés par Fagerlund *et al.* (2014), qui suggèrent que Spo0A n'est pas impliqué dans la mobilité par swimming chez *B. thuringiensis*. Cependant, les tests de mobilité utilisés dans l'étude sus-citée impliquent des cellules en phase exponentielle de croissance, dans lesquelles la concentration de Spo0A est très faible. Dans ces conditions, on peut donc imaginer que le mutant Δ*spo0A* se comportera comme la souche parentale, et sera notamment aussi mobile. Pour prouver une éventuelle augmentation de mobilité de la souche Δ*spo0A* en phase stationnaire de croissance, il faudrait plutôt entreprendre des expériences d'imagerie afin de comparer la mobilité d'un état frais et le nombre de flagelles (par coloration) du mutant et de la souche parentale, sur des cellules échantillonnées au cours de cette phase.

On note que le régulon Spo0A central contient également le gène BTB_c45020. BTB_c45020 code pour un régulateur dont le motif de liaison à l'ADN est semblable à celui de SinR, et qui a été récemment caractérisé dans l'équipe comme étant impliqué dans la mobilité (M. Gohar, communication personnelle). Cependant, d'après le modèle actuel, BTB_c45020 serait stimulateur de la mobilité, et cela ne permet donc pas d'expliquer la sur-expression des gènes de mobilité dans le mutant Δ*spo0A*. Il faut donc supposer que l'implication de Spo0A dans la mobilité est complexe et implique plusieurs voies.

2.4 Les potentiels régulateurs transcriptionnels sous le contrôle de Spo0A

Le régulon Spo0A central comprend 17 gènes dont le produit prédit est un régulateur transcriptionnel. Parmi ces 17 régulateurs potentiels, 8 n'ont aucune autre annotation que « régulateur transcriptionnel ». Les 9 restants sont PlcR, AbrB1, MogR, BTB_c45020, TreR, GlcT, CspC, MhqR2 et RocR. Chez *B. subtilis*, TreR réprime la transcription de l'opéron tréhalose (Schöck and Dahl, 1996), GlcT réprime l'expression de l'opéron *ptsGHI* (Stülke *et al.*, 1997), MhqR est impliqué dans la résistance à des molécules toxiques (Töwe *et al.*, 2007), RocR permet le catabolisme de l'arginine (Calogero *et al.*, 1994), et l'expression de *cspC* est induite par un choc thermique froid (Graumann *et al.*, 1996).

Il serait intéressant, dans un premier temps, de tenter de caractériser les régulateurs sans fonction attribuée. Définir, dans un second temps, le régulon de certains de ces régulateurs pourrait être envisagé, à commencer éventuellement par AbrB1, MogR et BTB_c45020.

2.5 Les plasmides

Les plasmides du groupe *B. cereus* sont probablement un de ses principaux vecteurs d'adaptation de ses différentes espèces à leurs niches écologiques respectives. On notera par exemple les plasmides pXO1 et pXO2 de *B. anthracis*, le plasmide portant le locus *ces* des souches émétiques de *B. cereus*, et les plasmides Cry de *B. thuringiensis*.

Plusieurs gènes plasmidiques font partie du régulon Spo0A central de *B. thuringiensis*. Cependant, le mapping des lectures RNA-seq des plasmides montre, contrairement à celles du chromosome, une assez mauvaise corrélation avec les annotations de gènes du génome de référence. Ces données sont donc à relativiser, car les gènes annotés n'en sont pas forcément, et que les boîtes Spo0A prédites dans des régions 5' non-transcrites sont en fait parfois retrouvées dans les données RNA-seq, donc transcrites. Il semble donc nécessaire, avant tout, de ré-annoter proprement les plasmides, notamment à partir des lectures RNA-seq.

Indépendamment de ces problèmes d'annotation, nous avons noté que tous les gènes d'un petit plasmide de 9 kb, pBTB_9p, étaient sous-exprimés avec des facteurs très importants dans la souche Δ*spo0A* comparée à la souche parentale. Notre hypothèse actuelle est que la réplication de ce plasmide nécessite Spo0A, et des expériences de qPCR sur les

deux souches permettraient de le vérifier.

3 Caractérisation des effecteurs supposés importants pour le nécrotrophisme

Nous avons pu, grâce à la comparaison entre les données acquises *in vitro* et *in vivo*, proposer un ensemble de 40 gènes du régulon Spo0A central potentiellement importants pour le nécrotrophisme de *B. thuringiensis* dans *Galleria mellonella*. Ces 40 gènes sont impliqués dans des fonctions très diverses et sont en cours d'analyse. Pour valider notre approche et éventuellement identifier des effecteurs importants du nécrotrophisme, il serait intéressant de construire les mutants de délétion de plusieurs de ces gènes et d'observer leur phénotype au cours de l'infection.

VALORISATION DE LA THÈSE

- **Communications orales et affichées**
  - Présentation orale « Genetic regulation of kurstakin, a biosurfactant produced by *Bacillus thuringiensis* », S. Gélis-Jeanvoine, M. Gohar, M. Gominet, L. Slamti, D. Lereclus. À l'occasion du colloque « Bactéries sporulantes pathogènes ou d'intérêt technologique » (Paris, France). Prix de la meilleure communication.
  - Poster « Transcription and role of kurstakin, a lipopeptide produced by *Bacillus thuringiensis* », S. Gélis-Jeanvoine, T. Caradec, S. Perchat, M. Gohar, M. Gominet, P. Jacques, D. Lereclus, L. Slamti. À l'occasion de la 8ème conférence internationale sur les micro-organismes à Gram positif (Montecatini Terme, Italie).

- **Articles dans des journaux revus par des pairs**
  - « Transcription and role of kurstakin, a lipopeptide produced by *Bacillus thuringiensis* », S. Gélis-Jeanvoine, A. Canette, M. Gohar, M. Gominet, L. Slamti, D. Lereclus. Soumis à Research in Microbiology.
  - « The Spo0A regulon of *Bacillus thuringiensis* is required for its necrotrophism », S. Gélis-Jeanvoine, D. Lereclus, L. Slamti. En préparation.

- **Brevet**
  - « Souche bactérienne génétiquement modifiée produisant de la kurstakine dans le milieu de culture », Didier Lereclus, Philippe Jacques, Michel Gohar, Marlène Chollet, Thibault Caradec, Sébastien Gélis-Jeanvoine. Déposé en France le 22/01/16.

- **Logiciels et scripts développés**
  - BLASTats, https://github.com/seb-ksl/BLASTats/
  - EggnogGenome, https://github.com/seb-ksl/EggnogGenome/
  - PyStrains, https://github.com/seb-ksl/PyStrains/

RÉFÉRENCES BIBLIOGRAPHIQUES

Abderrahmani, A., Tapi, A., Nateche, F., Chollet, M., Leclère, V., Wathelet, B., Hacene, H., and Jacques, P. (2011). Bioinformatics and molecular approaches to detect NRPS genes involved in the biosynthesis of kurstakin from Bacillus thuringiensis. Appl. Microbiol. Biotechnol. *92*, 571–581.

Abi Khattar, Z., Rejasse, A., Destoumieux-Garzón, D., Escoubas, J.M., Sanchis, V., Lereclus, D., Givaudan, A., Kallassy, M., Nielsen-Leroux, C., and Gaudriault, S. (2009). The dlt operon of Bacillus cereus is required for resistance to cationic antimicrobial peptides and for virulence in insects. J. Bacteriol. *191*, 7063–7073.

Agaisse, H., and Lereclus, D. (1994). Expression in Bacillus subtilis of the Bacillus thuringiensis cryIIIA toxin gene is not dependent on a sporulation-specific sigma factor and is increased in a spo0A mutant. J. Bacteriol. *176*, 4734–4741.

Agaisse, H., Gominet, M., Okstad, O.A., Kolstø, A.B., and Lereclus, D. (1999). PlcR is a pleiotropic regulator of extracellular virulence factor gene expression in Bacillus thuringiensis. Mol. Microbiol. *32*, 1043–1053.

Agata, N., Mori, M., Ohta, M., Suwan, S., Ohtani, I., and Isobe, M. (1994). A novel dodecadepsipeptide, cereulide, isolated from Bacillus cereus causes vacuole formation in HEp-2 cells. FEMS Microbiol. Lett. *121*, 31–34.

Aguilar, C., Vlamakis, H., Guzman, A., Losick, R., and Kolter, R. (2010). KinD is a checkpoint protein linking spore formation to extracellular-matrix production in Bacillus subtilis biofilms. MBio *1*.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389–3402.

Andersson, D.I., and Hughes, D. (2014). Microbiological effects of sublethal levels of antibiotics. Nat. Rev. Microbiol. *12*, 465–478.

Angus, T.A. (1953). Studies of Bacillus spp. pathogenic for silkworm. Bi-Monthly Prog. Report, Canada Dep. Sci. Serv. For. Biol. Div. *9*, 1–2.

Angus, T.A. (1954). A Bacterial Toxin paralysing Silkworm Larvæ. Nature *173*, 545–546.

ANSES (2011). Fiche de description de danger biologique transmissible par les aliments : “Bacillus cereus”-septembre 2011.

Aoki, K., and Chigasaki, Y. (1915). Über die Pathogenität der sog. Sotto-Bacillen (Ishiwata) bei Seidenraupen. Mitteilungen Aus Der Midizinischen Fak. Der Kais. Univ. Zu Tokyo *13*, 419–440.

Arima, K., Kakinuma, A., and Tamura, G. (1968). Surfactin, a crystalline peptidelipid surfactant produced by Bacillus subtilis: isolation, characterization and its inhibition of fibrin clot formation. Biochem. Biophys. Res. Commun. *31*, 488–494.

Awad, M.M., Johanesen, P.A., Carter, G.P., Rose, E., and Lyras, D. (2014). Clostridium difficile virulence factors: Insights into an anaerobic spore-forming pathogen. Gut Microbes *5*, 579–593.

Bahl, M.I., and Rosenberg, K. (2010). High abundance and diversity of Bacillus anthracis plasmid pXO1-like replicons in municipal wastewater. FEMS Microbiol. Ecol. *74*, 241–247.

Bai, U., Mandic-Mulec, I., and Smith, I. (1993). SinI modulates the activity of SinR, a developmental switch protein of Bacillus subtilis, by protein-protein interaction. Genes Dev. *7*, 139–148.

Baldus, J.M., Green, B.D., Youngman, P., and Moran, C.P. (1994). Phosphorylation of Bacillus subtilis transcription factor Spo0A stimulates transcription from the spoIIG promoter by enhancing binding to weak 0A boxes. J. Bacteriol. *176*, 296–306.

Béchet, M., Caradec, T., Hussein, W., Abderrahmani, A., Chollet, M., Leclère, V., Dubois, T., Lereclus, D., Pupin, M., and Jacques, P. (2012). Structure, biosynthesis, and properties of kurstakins, nonribosomal lipopeptides from Bacillus spp. Appl. Microbiol. Biotechnol. *95*, 593–600.

Belitsky, B.R., and Sonenshein, A.L. (2013). Genome-wide identification of Bacillus subtilis CodY-binding sites at single-nucleotide resolution. Proc. Natl. Acad. Sci. U. S. A. *110*, 7026–7031.

Bennaceur, I. (2013). Etude du recrutement de la phase planctonique par le biofilm chez Bacillus cereus : approches physiologiques et moléculaires. Paris 11.

Berliner, E. (1911). Über die Schlafsucht der Mehlmottenraupe. Z. Gesamte Getreidewes. *3*, 63.

Berliner, E. (1915). Über die Schlafsucht der Mehlmottenraupe (Ephestia kühniella Zell.) und ihren Erreger Bacillus thuringiensis n. sp. Z. Angew. Entomol. *2*, 29–56.

Boguslawski, K.M., Hill, P.A., and Griffith, K.L. (2015). Novel mechanisms of controlling the activities of the transcription factors Spo0A and ComA by the plasmid-encoded quorum sensing regulators Rap60-Phr60 in Bacillus subtilis. Mol. Microbiol. *96*, 325–348.

Branda, S.S., Chu, F., Kearns, D.B., Losick, R., and Kolter, R. (2006). A major protein component of the Bacillus subtilis biofilm matrix. Mol. Microbiol. *59*, 1229–1238.

Brézillon, C., Haustant, M., Dupke, S., Corre, J.-P., Lander, A., Franz, T., Monot, M., Couture-Tosi, E., Jouvion, G., Leendertz, F.H., *et al.* (2015). Capsules, toxins and AtxA as virulence factors of emerging Bacillus cereus biovar anthracis. PLoS Negl. Trop. Dis. *9*, e0003455.

Brown, L., Kessler, A., Cabezas-Sanchez, P., Luque-Garcia, J.L., and Casadevall, A. (2014). Extracellular Vesicles Produced by the Gram-positive Bacterium Bacillus subtilis are Disrupted by the Lipopeptide Surfactin. Mol. Microbiol. n/a – n/a.

Bumpus, S.B., Evans, B.S., Thomas, P.M., Ntai, I., and Kelleher, N.L. (2009). A proteomics approach to discovering natural products and their biosynthetic pathways. Nat. Biotechnol. *27*, 951–956.

Burbulys, D., Trach, K.A., and Hoch, J.A. (1991). Initiation of sporulation in B. subtilis is controlled by a multicomponent phosphorelay. Cell *64*, 545–552.

Burkholder, W.F., Kurtser, I., and Grossman, A.D. (2001). Replication initiation proteins regulate a developmental checkpoint in Bacillus subtilis. Cell *104*, 269–279.

Calogero, S., Gardan, R., Glaser, P., Schweizer, J., Rapoport, G., and Debarbouille, M. (1994). RocR, a novel regulatory protein controlling arginine utilization in Bacillus subtilis, belongs to the NtrC/NifA family of transcriptional activators. J. Bacteriol. *176*, 1234–1241.

Cameotra, S.S., and Makkar, R.S. (2004). Recent applications of biosurfactants as biological and immunological molecules. Curr. Opin. Microbiol. *7*, 262–266.

Cawoy, H., Mariutto, M., Henry, G., Fisher, C., Vasilyeva, N., Thonart, P., Dommes, J., and Ongena, M. (2013). Plant defense stimulation by natural isolates of Bacillus depends on efficient surfactin production. Mol. Plant. Microbe. Interact.

Chai, Y., Chu, F., Kolter, R., and Losick, R. (2008). Bistability and biofilm formation in Bacillus subtilis. Mol. Microbiol. *67*, 254–263.

Chastanet, A., and Losick, R. (2011). Just-in-time control of Spo0A synthesis in Bacillus subtilis by multiple regulatory mechanisms. J. Bacteriol. *193*, 6366–6374.

Chastanet, A., Vitkup, D., Yuan, G.-C., Norman, T.M., Liu, J.S., and Losick, R.M. (2010). Broadly heterogeneous activation of the master regulator for sporulation in Bacillus subtilis. Proc. Natl. Acad. Sci. U. S. A. *107*, 8486–8491.

Château, A., van Schaik, W., Joseph, P., Handke, L.D., McBride, S.M., Smeets, F.M.H., Sonenshein, A.L., and Fouet, A. (2013). Identification of CodY targets in Bacillus anthracis by genome-wide in vitro binding analysis. J. Bacteriol. *195*, 1204–1213.

Chen, H.-J., Tsai, T.-K., Pan, S.-C., Lin, J.-S., Tseng, C.-L., and Shaw, G.-C. (2010). The master transcription factor Spo0A is required for poly(3-hydroxybutyrate) (PHB) accumulation and expression of genes involved in PHB biosynthesis in Bacillus thuringiensis. FEMS Microbiol. Lett. *304*, 74–81.

Chibazakura, T., Kawamura, F., and Takahashi, H. (1991). Differential regulation of spo0A transcription in Bacillus subtilis: glucose represses promoter switching at the initiation of sporulation. J. Bacteriol. *173*, 2625–2632.

Chu, F., Kearns, D.B., Branda, S.S., Kolter, R., and Losick, R. (2006). Targets of the master regulator of biofilm formation in Bacillus subtilis. Mol. Microbiol. *59*, 1216–1228.

Chu, F., Kearns, D.B., McLoon, A., Chai, Y., Kolter, R., and Losick, R. (2008). A novel regulatory protein governing biofilm formation in Bacillus subtilis. Mol. Microbiol. *68*, 1117–

1127.

Clark, T.M. (1999). Evolution and Adaptive Significance of Larval Midgut Alkalinization in the Insect Superorder Mecopterida. J. Chem. Ecol. *25*, 1945–1960.

Cock, P.J.A., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., *et al.* (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics *25*, 1422–1423.

Dalhammar, G., and Steiner, H. (1984). Characterization of inhibitor A, a protease from Bacillus thuringiensis which degrades attacins and cecropins, two classes of antibacterial proteins in insects. Eur. J. Biochem. *139*, 247–252.

Đapa, T., Leuzzi, R., Ng, Y.K., Baban, S.T., Adamo, R., Kuehne, S.A., Scarselli, M., Minton, N.P., Serruto, D., and Unnikrishnan, M. (2013). Multiple factors modulate biofilm formation by the anaerobic pathogen Clostridium difficile. J. Bacteriol. *195*, 545–555.

Dartois, V., Djavakhishvili, T., and Hoch, J.A. (1996). Identification of a membrane protein involved in activation of the KinB pathway to sporulation in Bacillus subtilis. J. Bacteriol. *178*, 1178–1186.

Dartois, V., Djavakhishvili, T., and Hoch, J.A. (1997). KapB is a lipoprotein required for KinB signal transduction and activation of the phosphorelay to sporulation in Bacillus subtilis. Mol. Microbiol. *26*, 1097–1108.

Dawson, L.F., Valiente, E., Faulds-Pain, A., Donahue, E.H., and Wren, B.W. (2012). Characterisation of Clostridium difficile biofilm formation, a role for Spo0A. PLoS One *7*, e50527.

Declerck, N., Bouillaut, L., Chaix, D., Rugani, N., Slamti, L., Hoh, F., Lereclus, D., and Arold, S.T. (2007). Structure of PlcR: Insights into virulence regulation and evolution of quorum sensing in Gram-positive bacteria. Proc. Natl. Acad. Sci. U. S. A. *104*, 18490–18495.

Delcambe, L. (1950). Iturine, new antibiotic produced by Bacillus subtilis. C. R. Seances Soc. Biol. Fil. *144*, 1431–1434.

Deleu, M., Lorent, J., Lins, L., Brasseur, R., Braun, N., El Kirat, K., Nylander, T., Dufrêne, Y.F., and Mingeot-Leclercq, M.-P. (2013). Effects of surfactin on membrane models displaying lipid phase separation. Biochim. Biophys. Acta *1828*, 801–815.

Deng, C., Peng, Q., Song, F., and Lereclus, D. (2014). Regulation of cry gene expression in Bacillus thuringiensis. Toxins (Basel). *6*, 2194–2209.

Dubois, T., Faegri, K., Perchat, S., Lemy, C., Buisson, C., Nielsen-LeRoux, C., Gohar, M., Jacques, P., Ramarao, N., Kolstø, A.-B., *et al.* (2012). Necrotrophism is a quorum-sensing-regulated lifestyle in Bacillus thuringiensis. PLoS Pathog. *8*, e1002629.

Dubois, T., Perchat, S., Verplaetse, E., Gominet, M., Lemy, C., Aumont-Nicaise, M., Grenha, R., Nessler, S., and Lereclus, D. (2013). Activity of the Bacillus thuringiensis NprR-NprX

cell-cell communication system is co-ordinated to the physiological stage through a complex transcriptional regulation. Mol. Microbiol. *88*, 48–63.

Edwards, A.N., and McBride, S.M. (2014). Initiation of sporulation in Clostridium difficile : a twist on the classic model. FEMS Microbiol. Lett. *358*, 110–118.

Ehling-Schulz, M., Fricker, M., Grallert, H., Rieck, P., Wagner, M., and Scherer, S. (2006). Cereulide synthetase gene cluster from emetic Bacillus cereus: structure and location on a mega virulence plasmid related to Bacillus anthracis toxin plasmid pXO1. BMC Microbiol. *6*, 20.

Eijlander, R.T., de Jong, A., Krawczyk, A.O., Holsappel, S., and Kuipers, O.P. (2014). SporeWeb: an interactive journey through the complete sporulation cycle of Bacillus subtilis. Nucleic Acids Res. *42*, D685–D691.

Ellermeier, C.D., Hobbs, E.C., Gonzalez-Pastor, J.E., and Losick, R. (2006). A three-protein signaling pathway governing immunity to a bacterial cannibalism toxin. Cell *124*, 549–559.

Fagerlund, A., Dubois, T., Økstad, O.-A., Verplaetse, E., Gilois, N., Bennaceur, I., Perchat, S., Gominet, M., Aymerich, S., Kolstø, A.-B., *et al.* (2014). SinR Controls Enterotoxin Expression in Bacillus thuringiensis Biofilms. PLoS One *9*, e87532.

Fajardo, A., and Martínez, J.L. (2008). Antibiotics as signals that trigger specific bacterial responses. Curr. Opin. Microbiol. *11*, 161–167.

Finking, R., and Marahiel, M.A. (2004). Biosynthesis of nonribosomal peptides. Annu. Rev. Microbiol. *58*, 453–488.

Flügge, C. (1886). Die Mikroorganismen. F. C. W. Vogel, Leipzig.

van Frankenhuyzen, K. (2009). Insecticidal activity of Bacillus thuringiensis crystal proteins. J. Invertebr. Pathol. *101*, 1–16.

Frenzel, E., Doll, V., Pauthner, M., Lücking, G., Scherer, S., and Ehling-Schulz, M. (2012). CodY orchestrates the expression of virulence determinants in emetic Bacillus cereus by impacting key regulatory circuits. Mol. Microbiol. *85*, 67–88.

Fujita, M., and Sadaie, Y. (1998). Feedback loops involving Spo0A and AbrB in in vitro transcription of the genes involved in the initiation of sporulation in Bacillus subtilis. J. Biochem. *124*, 98–104.

Fujita, M., González-Pastor, J.E., and Losick, R. (2005). High- and low-threshold genes in the Spo0A regulon of Bacillus subtilis. J. Bacteriol. *187*, 1357–1368.

Gallego del Sol, F., and Marina, A. (2013). Structural basis of Rap phosphatase inhibition by Phr peptides. PLoS Biol. *11*, e1001511.

Galperin, M.Y. (2013). Genome Diversity of Spore-Forming Firmicutes. Microbiol. Spectr. *1*, TBS0015–TBS2012.

Garti-Levi, S., Eswara, A., Smith, Y., Fujita, M., and Ben-Yehuda, S. (2013). Novel modulators controlling entry into sporulation in Bacillus subtilis. J. Bacteriol. *195*, 1475–1483.

Gohar, M., Økstad, O.A., Gilois, N., Sanchis, V., Kolstø, A.-B., and Lereclus, D. (2002). Two-dimensional electrophoresis analysis of the extracellular proteome of Bacillus cereus reveals the importance of the PlcR regulon. Proteomics *2*, 784–791.

Gohar, M., Faegri, K., Perchat, S., Ravnum, S., Økstad, O.A., Gominet, M., Kolstø, A.-B., and Lereclus, D. (2008). The PlcR virulence regulon of Bacillus cereus. PLoS One *3*, e2793.

González-Pastor, J.E., Hobbs, E.C., and Losick, R. (2003). Cannibalism by sporulating bacteria. Science *301*, 510–513.

Grau, R.R., de Oña, P., Kunert, M., Leñini, C., Gallegos-Monterrosa, R., Mhatre, E., Vileta, D., Donato, V., Hölscher, T., Boland, W., *et al.* (2015). A Duo of Potassium-Responsive Histidine Kinases Govern the Multicellular Destiny of Bacillus subtilis. MBio *6*.

Graumann, P., Schröder, K., Schmid, R., and Marahiel, M.A. (1996). Cold shock stress-induced proteins in Bacillus subtilis. J. Bacteriol. *178*, 4611–4619.

Grenha, R., Slamti, L., Nicaise, M., Refes, Y., Lereclus, D., and Nessler, S. (2012). Structural basis for the activation mechanism of the PlcR virulence regulator by the quorum-sensing signal peptide PapR. Proc. Natl. Acad. Sci. *110*, 1047–1052.

Guillemet, E., Cadot, C., Tran, S.-L., Guinebretière, M.-H., Lereclus, D., and Ramarao, N. (2010). The InhA metalloproteases of Bacillus cereus contribute concomitantly to virulence. J. Bacteriol. *192*, 286–294.

Guinebretière, M.-H., Thompson, F.L., Sorokin, A., Normand, P., Dawyndt, P., Ehling-Schulz, M., Svensson, B., Sanchis, V., Nguyen-The, C., Heyndrickx, M., *et al.* (2008). Ecological diversification in the Bacillus cereus Group. Environ. Microbiol. *10*, 851–865.

Haldenwang, W.G. (1995). The sigma factors of Bacillus subtilis. Microbiol. Rev. *59*, 1–30.

Hamon, M.A., Stanley, N.R., Britton, R.A., Grossman, A.D., and Lazazzera, B.A. (2004). Identification of AbrB-regulated genes involved in biofilm formation by Bacillus subtilis. Mol. Microbiol. *52*, 847–860.

Hannay, C.L. (1953). Crystalline inclusions in aerobic spore-forming bacteria. Nature *172*, 1004.

Hathout, Y., Ho, Y.P., Ryzhov, V., Demirev, P., and Fenselau, C. (2000). Kurstakins: a new class of lipopeptides isolated from Bacillus thuringiensis. J. Nat. Prod. *63*, 1492–1496.

Heimpel, A.M., and Angus, T.A. (1959). The site of action of crystalliferous bacteria in Lepidoptera larvae. J. Invertebr. Pathol. *1*, 152–170.

Higgins, D., and Dworkin, J. (2012). Recent progress in Bacillus subtilis sporulation. FEMS

Microbiol. Rev. *36*, 131–148.

Hu, X., Van der Auwera, G., Timmery, S., Zhu, L., and Mahillon, J. (2009). Distribution, diversity, and potential mobility of extrachromosomal elements related to the Bacillus anthracis pXO1 and pXO2 virulence plasmids. Appl. Environ. Microbiol. *75*, 3016–3028.

Huerta-Cepas, J., Dopazo, J., and Gabaldón, T. (2010). ETE: a python Environment for Tree Exploration. BMC Bioinformatics *11*, 24.

Huillet, E., Tempelaars, M.H., André-Leroux, G., Wanapaisan, P., Bridoux, L., Makhzami, S., Panbangred, W., Martin-Verstraete, I., Abee, T., and Lereclus, D. (2012). PlcRa, a new quorum-sensing regulator from Bacillus cereus, plays a role in oxidative stress responses and cysteine metabolism in stationary phase. PLoS One *7*, e51047.

Iatsenko, I., Yim, J.J., Schroeder, F.C., and Sommer, R.J. (2014). B. subtilis GS67 protects C. elegans from Gram-positive pathogens via fengycin-mediated microbial antagonism. Curr. Biol. *24*, 2720–2727.

Inès, M., and Dhouha, G. (2015). Lipopeptide surfactants: Production, Recovery and Pore Forming Capacity. Peptides *71*, 100–112.

Ishii, T. (2001). DBTBS: a database of Bacillus subtilis promoters and transcription factors. Nucleic Acids Res. *29*, 278–280.

Ivanova, N., Sorokin, A., Anderson, I., Galleron, N., Candelon, B., Kapatral, V., Bhattacharyya, A., Reznik, G., Mikhailova, N., Lapidus, A., *et al.* (2003). Genome sequence of Bacillus cereus and comparative analysis with Bacillus anthracis. Nature *423*, 87–91.

Jacques, P. (2011). Surfactin and other lipopeptides from Bacillus spp. In Biosurfactants, G. Soberón-Chávez, ed. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 57–91.

Jarrell, K.F., and McBride, M.J. (2008). The surprisingly diverse ways that prokaryotes move. Nat. Rev. Microbiol. *6*, 466–476.

Jiang, M., Shao, W., Perego, M., and Hoch, J.A. (2000). Multiple histidine kinases regulate entry into stationary phase and sporulation in Bacillus subtilis. Mol. Microbiol. *38*, 535–542.

Jiménez, G., Urdiain, M., Cifuentes, A., López-López, A., Blanch, A.R., Tamames, J., Kämpfer, P., Kolstø, A.-B., Ramón, D., Martínez, J.F., *et al.* (2013). Description of Bacillus toyonensis sp. nov., a novel species of the Bacillus cereus group, and pairwise genome comparisons of the species of the group by means of ANI calculations. Syst. Appl. Microbiol. *36*, 383–391.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., *et al.* (2014). InterProScan 5: genome-scale protein function classification. Bioinformatics *30*, 1236–1240.

Julkowska, D., Obuchowski, M., Holland, I.B., and Séror, S.J. (2005). Comparative analysis of the development of swarming communities of Bacillus subtilis 168 and a natural wild type:

critical effects of surfactin and the composition of the medium. J. Bacteriol. *187*, 65–76.

Kamar, R. (2014). Mécanismes de résistance aux peptides antimicrobiens chez Bacillus thuringiensis : rôle de dltX dans la D-alanylation des acides téichoïques. AgroParisTech.

Kanlayavattanakul, M., and Lourith, N. (2010). Lipopeptides in cosmetics. Int. J. Cosmet. Sci. *32*, 1–8.

Kearns, D.B., Chu, F., Branda, S.S., Kolter, R., and Losick, R. (2005). A master regulator for biofilm formation by Bacillus subtilis. Mol. Microbiol. *55*, 739–749.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. *10*, 845–858.

Khong, N.G., Randoux, B., Tayeh, C., Coutte, F., Bourdon, N., Tisserant, B., Laruelle, F., Jacques, P., and Reignault, P. (2012). Induction of resistance in wheat against powdery mildew by bacterial cyclic lipopeptides. Commun. Agric. Appl. Biol. Sci. *77*, 39–51.

Khong, N.G., Randoux, B., Deravel, J., Tisserant, B., Tayeh, C., Coutte, F., Bourdon, N., Jacques, P., and Reignault, P. (2013). Induction of resistance in wheat by bacterial cyclic lipopeptides. Commun. Agric. Appl. Biol. Sci. *78*, 479–487.

Kobayashi, K. (2007). Bacillus subtilis pellicle formation proceeds through genetically defined morphological changes. J. Bacteriol. *189*, 4920–4931.

Koch, R. (1876). Die Ätiologie der Milzbrand-Krankheit, begründet auf die Entwicklungsgeschichte des Bacillus Anthracis. Beiträge Zur Biol. Der Pflanz. *2*.

Kovács, Á.T. (2015). Bacterial differentiation via gradual activation of global regulators. Curr. Genet.

de la Fuente-Salcido, N.M., Casados-Vázquez, L.E., and Barboza-Corona, J.E. (2013). Bacteriocins of Bacillus thuringiensis can expand the potential of this bacterium to other areas rather than limit its use only as microbial insecticide. Can. J. Microbiol. *59*, 515–522.

Lapidus, A., Goltsman, E., Auger, S., Galleron, N., Ségurens, B., Dossat, C., Land, M.L., Broussolle, V., Brillard, J., Guinebretiere, M.-H., *et al.* (2008). Extending the Bacillus cereus group genomics to putative food-borne pathogens of different toxicity. Chem. Biol. Interact. *171*, 236–249.

Lechner, S., Mayr, R., Francis, K.P., Prüss, B.M., Kaplan, T., Wiessner-Gunkel, E., Stewart, G.S., and Scherer, S. (1998). Bacillus weihenstephanensis sp. nov. is a new psychrotolerant species of the Bacillus cereus group. Int. J. Syst. Bacteriol. *48 Pt 4*, 1373–1382.

Lereclus, D., Agaisse, H., Gominet, M., and Chaufaux, J. (1995). Overproduction of Encapsulated Insecticidal Crystal Proteins in a Bacillus thuringiensis spoOA Mutant. Bio/Technology *13*, 67–71.

Lereclus, D., Agaisse, H., Gominet, M., Salamitou, S., and Sanchis, V. (1996). Identification

of a Bacillus thuringiensis gene that positively regulates transcription of the phosphatidylinositol-specific phospholipase C gene at the onset of the stationary phase. J. Bacteriol. *178*, 2749–2756.

Lereclus, D., Agaisse, H., Grandvalet, C., Salamitou, S., and Gominet, M. (2000). Regulation of toxin and virulence gene transcription in Bacillus thuringiensis. Int. J. Med. Microbiol. *290*, 295–299.

Lewis, R.J., Muchová, K., Brannigan, J.A., Barák, I., Leonard, G., and Wilkinson, A.J. (2000). Domain swapping in the sporulation response regulator Spo0A. J. Mol. Biol. *297*, 757–770.

Li, R.S., Jarrett, P., and Burges, H.D. (1987). Importance of spores, crystals, and δ-endotoxins in the pathogenicity of different varieties of Bacillus thuringiensis in Galleria mellonella and Pieris brassicae. J. Invertebr. Pathol. *50*, 277–284.

Li, X., Yang, H., Zhang, D., Li, X., Yu, H., and Shen, Z. (2015). Overexpression of specific proton motive force-dependent transporters facilitate the export of surfactin in Bacillus subtilis. J. Ind. Microbiol. Biotechnol. *42*, 93–103.

Lindbäck, T., Mols, M., Basset, C., Granum, P.E., Kuipers, O.P., and Kovács, Á.T. (2012). CodY, a pleiotropic regulator, influences multicellular behaviour and efficient production of virulence factors in Bacillus cereus. Environ. Microbiol. *14*, 2233–2246.

Little, T.J., Hultmark, D., and Read, A.F. (2005). Invertebrate immunity and the limits of mechanistic immunology. Nat. Immunol. *6*, 651–654.

López, D., and Kolter, R. (2010). Extracellular signals that define distinct and coexisting cell fates in Bacillus subtilis. FEMS Microbiol. Rev. *34*, 134–149.

López, D., Vlamakis, H., and Kolter, R. (2009a). Generation of multiple cell types in Bacillus subtilis. FEMS Microbiol. Rev. *33*, 152–163.

López, D., Fischbach, M.A., Chu, F., Losick, R., and Kolter, R. (2009b). Structurally diverse natural products that cause potassium leakage trigger multicellularity in Bacillus subtilis. Proc. Natl. Acad. Sci. U. S. A. *106*, 280–285.

López, D., Vlamakis, H., Losick, R., and Kolter, R. (2009c). Cannibalism enhances biofilm development in Bacillus subtilis. Mol. Microbiol. *74*, 609–618.

López, D., Vlamakis, H., Losick, R., and Kolter, R. (2009d). Paracrine signaling in a bacterium. Genes Dev. *23*, 1631–1638.

Lozano Goné, A.M., Dinorín Téllez Girón, J., Jiménez Montejo, F.E., Hidalgo-Lara, M.E., and López y López, V.E. (2014). Behavior of Transition State Regulator AbrB in Batch Cultures of Bacillus thuringiensis. Curr. Microbiol. *69*, 725–732.

Lücking, G., Dommel, M.K., Scherer, S., Fouet, A., and Ehling-Schulz, M. (2009). Cereulide synthesis in emetic Bacillus cereus is controlled by the transition state regulator AbrB, but not

by the virulence regulator PlcR. Microbiology *155*, 922–931.

Mackin, K.E., Carter, G.P., Howarth, P., Rood, J.I., and Lyras, D. (2013). Spo0A differentially regulates toxin production in evolutionarily diverse strains of Clostridium difficile. PLoS One *8*, e79666.

Magnuson, R., Solomon, J., and Grossman, A.D. (1994). Biochemical and genetic characterization of a competence pheromone from B. subtilis. Cell *77*, 207–216.

Mani, N., and Dupuy, B. (2001). Regulation of toxin synthesis in Clostridium difficile by an alternative RNA polymerase sigma factor. Proc. Natl. Acad. Sci. U. S. A. *98*, 5844–5849.

Maughan, H., and Van der Auwera, G. (2011). Bacillus taxonomy in the genomic era finds phenotypes to be essential though often misleading. Infect. Genet. Evol. *11*, 789–797.

May, J.J., Wendrich, T.M., and Marahiel, M.A. (2000). The dhb Operon of Bacillus subtilisEncodes the Biosynthetic Template for the Catecholic Siderophore 2,3-Dihydroxybenzoate-Glycine-Threonine Trimeric Ester Bacillibactin. J. Biol. Chem. *276*, 7209–7217.

Mazzola, M., de Bruijn, I., Cohen, M.F., and Raaijmakers, J.M. (2009). Protozoan-induced regulation of cyclic lipopeptide biosynthesis is an effective predation defense mechanism for Pseudomonas fluorescens. Appl. Environ. Microbiol. *75*, 6804–6811.

Medema, M.H., Blin, K., Cimermancic, P., de Jager, V., Zakrzewski, P., Fischbach, M.A., Weber, T., Takano, E., and Breitling, R. (2011). antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. Nucleic Acids Res. *39*, W339–W346.

Mellin, J.R., and Cossart, P. (2015). Unexpected versatility in bacterial riboswitches. Trends Genet. *31*, 150–156.

Moïse (700BC). L’exode. In La Bible, (Vatican),.

Molle, V., Fujita, M., Jensen, S.T., Eichenberger, P., González-Pastor, J.E., Liu, J.S., and Losick, R. (2003a). The Spo0A regulon of Bacillus subtilis. Mol. Microbiol. *50*, 1683–1701.

Molle, V., Nakaura, Y., Shivers, R.P., Yamaguchi, H., Losick, R., Fujita, Y., and Sonenshein, A.L. (2003b). Additional Targets of the Bacillus subtilis Global Regulator CodY Identified by Chromatin Immunoprecipitation and Genome-Wide Transcript Analysis. J. Bacteriol. *185*, 1911–1922.

Moran, C.P.J. (1993). RNA polymerase and transcription factors. In Bacillus Subtilis and Other Gram-Positive Bacteria, A.L. Sonenshein, J.A. Hoch, and R. Losick, eds. (Washington D.C.: American Society for Microbiology), pp. 653–667.

Nakamura, L.K. (1998). Bacillus pseudomycoides sp. nov. Int. J. Syst. Bacteriol. *48 Pt 3*, 1031–1035.

Nasir, M.N., and Besson, F. (2012). Interactions of the antifungal mycosubtilin with ergosterol-containing interfacial monolayers. Biochim. Biophys. Acta *1818*, 1302–1308.

Newman, J.A., Rodrigues, C., and Lewis, R.J. (2013). Molecular Basis of the Activity of SinR Protein, the Master Regulator of Biofilm Formation in Bacillus subtilis. J. Biol. Chem. *288*, 10766–10778.

Nishikiori, T., Naganawa, H., Muraoka, Y., Aoyagi, T., and Umezawa, H. (1986). Plipastatins: new inhibitors of phospholipase A2, produced by Bacillus cereus BMG302-fF67. III. Structural elucidation of plipastatins. J. Antibiot. (Tokyo). *39*, 755–761.

Nitschke, M., and Costa, S.G.V.A.O. (2007). Biosurfactants in food industry. Trends Food Sci. Technol. *18*, 252–259.

Palma, L., Muñoz, D., Berry, C., Murillo, J., and Caballero, P. (2014). Bacillus thuringiensis toxins: an overview of their biocidal activity. Toxins (Basel). *6*, 3296–3325.

Parashar, V., Mirouze, N., Dubnau, D.A., and Neiditch, M.B. (2011). Structural basis of response regulator dephosphorylation by Rap phosphatases. PLoS Biol. *9*, e1000589.

Parashar, V., Jeffrey, P.D., and Neiditch, M.B. (2013a). Conformational change-induced repeat domain expansion regulates Rap phosphatase quorum-sensing signal receptors. PLoS Biol. *11*, e1001512.

Parashar, V., Konkol, M.A., Kearns, D.B., and Neiditch, M.B. (2013b). A plasmid-encoded phosphatase regulates Bacillus subtilis biofilm architecture, sporulation, and genetic competence. J. Bacteriol. *195*, 2437–2448.

Parker, G.F., Daniel, R.A., and Errington, J. (1996). Timing and genetic regulation of commitment to sporulation in Bacillus subtilis. Microbiology *142*, 3445–3452.

Perego, M. (2001). A new family of aspartyl phosphate phosphatases targeting the sporulation transcription factor Spo0A of Bacillus subtilis. Mol. Microbiol. *42*, 133–143.

Pettit, L.J., Browne, H.P., Yu, L., Smits, W.K., Fagan, R.P., Barquist, L., Martin, M.J., Goulding, D., Duncan, S.H., Flint, H.J., *et al.* (2014). Functional genomics reveals that Clostridium difficile Spo0A coordinates sporulation, virulence and metabolism. BMC Genomics *15*, 160.

Peypoux, F., Michel, G., and Delcambe, L. (1976). The structure of mycosubtilin, an antibiotic isolated from Bacillus subtilis. Eur. J. Biochem. *63*, 391–398.

Peypoux, F., Guinand, M., Michel, G., Delcambe, L., Das, B.C., and Lederer, E. (1978). Structure of iturine A, a peptidolipid antibiotic from Bacillus subtilis. Biochemistry *17*, 3992–3996.

Pflughoeft, K.J., Sumby, P., and Koehler, T.M. (2011). Bacillus anthracis sin locus and regulation of secreted proteases. J. Bacteriol. *193*, 631–639.

Phillips, Z.E. V, and Strauch, M.A. (2002). Bacillus subtilis sporulation and stationary phase gene expression. Cell. Mol. Life Sci. *59*, 392–402.

Piggot, P.J., and Coote, J.G. (1976). Genetic aspects of bacterial endospore formation. Bacteriol. Rev. *40*, 908–962.

Piggot, P.J., and Hilbert, D.W. (2004). Sporulation of Bacillus subtilis. Curr. Opin. Microbiol. *7*, 579–586.

Poncet, S., Dervyn, E., Klier, A., and Rapoport, G. (1997). Spo0A represses transcription of the cry toxin genes in Bacillus thuringiensis. Microbiology *143 (Pt 8*, 2743–2751.

Pottathil, M., and Lazazzera, B.A. (2003). The extracellular Phr peptide-Rap phosphatase signaling circuit of Bacillus subtilis. Front. Biosci. *8*, d32–d45.

Powell, S., Forslund, K., Szklarczyk, D., Trachana, K., Roth, A., Huerta-Cepas, J., Gabaldón, T., Rattei, T., Creevey, C., Kuhn, M., *et al.* (2014). eggNOG v4.0: nested orthology inference across 3686 organisms. Nucleic Acids Res. *42*, D231–D239.

Prasertphon, S., Areekul, P., and Tanada, Y. (1973). Sporulation of Bacillus thuringiensis in host cadavers. J. Invertebr. Pathol. *21*, 205–207.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol. Biol. Evol. *26*, 1641–1650.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2--approximately maximum-likelihood trees for large alignments. PLoS One *5*, e9490.

Raaijmakers, J.M., De Bruijn, I., Nybroe, O., and Ongena, M. (2010). Natural functions of lipopeptides from Bacillus and Pseudomonas: more than surfactants and antibiotics. FEMS Microbiol. Rev. *34*, 1037–1062.

Ramarao, N., Nielsen-Leroux, C., and Lereclus, D. (2012). The Insect Galleria mellonella as a Powerful Infection Model to Investigate Bacterial Pathogenesis. J. Vis. Exp. e4392.

Ratnayake-Lecamwasam, M., Serror, P., Wong, K.W., and Sonenshein, A.L. (2001). Bacillus subtilis CodY represses early-stationary-phase genes by sensing GTP levels. Genes Dev. *15*, 1093–1103.

Rocha-Estrada, J., Aceves-Diez, A.E., Guarneros, G., and de la Torre, M. (2010). The RNPP family of quorum-sensing proteins in Gram-positive bacteria. Appl. Microbiol. Biotechnol. *87*, 913–923.

Romero, D., Aguilar, C., Losick, R., and Kolter, R. (2010). Amyloid fibers provide structural integrity to Bacillus subtilis biofilms. Proc. Natl. Acad. Sci. U. S. A. *107*, 2230–2234.

Romero, D., Vlamakis, H., Losick, R., and Kolter, R. (2011). An accessory protein required for anchoring and assembly of amyloid fibres in B. subtilis biofilms. Mol. Microbiol. *80*,

1155–1168.

Rosenbusch, K.E., Bakker, D., Kuijper, E.J., and Smits, W.K. (2012). C. difficile 630Δerm Spo0A regulates sporulation, but does not contribute to toxin production, by direct high-affinity binding to target DNA. PLoS One *7*, e48608.

Rost, B., Liu, J., Nair, R., Wrzeszczynski, K.O., and Ofran, Y. (2003). Automatic prediction of protein function. Cell. Mol. Life Sci. *60*, 2637–2650.

Rowe-Magnus, D.A., and Spiegelman, G.B. (1998). DNA strand separation during activation of a developmental promoter by the Bacillus subtilis response regulator Spo0A. Proc. Natl. Acad. Sci. U. S. A. *95*, 5305–5310.

Rowland, S.L., Burkholder, W.F., Cunningham, K.A., Maciejewski, M.W., Grossman, A.D., and King, G.F. (2004). Structure and mechanism of action of Sda, an inhibitor of the histidine kinases that regulate initiation of sporulation in Bacillus subtilis. Mol. Cell *13*, 689–701.

Rubinstein, S.M., Kolodkin-Gal, I., McLoon, A., Chai, L., Kolter, R., Losick, R., and Weitz, D.A. (2012). Osmotic pressure can regulate matrix gene expression in Bacillus subtilis. Mol. Microbiol. *86*, 426–436.

Salamitou, S., Ramisse, F., Brehelin, M., Bourguet, D., Gilois, N., Gominet, M., Hernandez, E., and Lereclus, D. (2000). The plcR regulon is involved in the opportunistic properties of Bacillus thuringiensis and Bacillus cereus in mice and insects. Microbiology *146*, 2825–2832.

Saujet, L., Pereira, F.C., Henriques, A.O., and Martin-Verstraete, I. (2014). The regulatory network controlling spore formation in Clostridium difficile. FEMS Microbiol. Lett. *358*, 1–10.

Schmidt, T.R., Scott, E.J., and Dyer, D.W. (2011). Whole-genome phylogenies of the family Bacillaceae and expansion of the sigma factor gene family in the Bacillus cereus species-group. BMC Genomics *12*, 430.

Schneider, T., Müller, A., Miess, H., and Gross, H. (2014). Cyclic lipopeptides as antibacterial agents - potent antibiotic activity mediated by intriguing mode of actions. Int. J. Med. Microbiol. *304*, 37–43.

Schöck, F., and Dahl, M.K. (1996). Expression of the tre operon of Bacillus subtilis 168 is regulated by the repressor TreR. J. Bacteriol. *178*, 4576–4581.

Schwarzer, D., Mootz, H.D., Linne, U., and Marahiel, M.A. (2002). Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases. Proc. Natl. Acad. Sci. U. S. A. *99*, 14083–14088.

Segond, D., Abi Khalil, E., Buisson, C., Daou, N., Kallassy, M., Lereclus, D., Arosio, P., Bou-Abdallah, F., and Nielsen Le Roux, C. (2014). Iron acquisition in Bacillus cereus: the roles of IlsA and bacillibactin in exogenous ferritin iron mobilization. PLoS Pathog. *10*, e1003935.

Shafikhani, S.H., Mandic-Mulec, I., Strauch, M.A., Smith, I., and Leighton, T. (2002).

Postexponential regulation of sin operon expression in Bacillus subtilis. J. Bacteriol. *184*, 564–571.

Shapiro, J.A. (1988). Bacteria as multicellular organisms. Sci. Am. 82–89.

Shapiro, J.A. (1998). Thinking about bacterial populations as multicellular organisms. Annu. Rev. Microbiol. *52*, 81–104.

Shemesh, M., and Chai, Y. (2013). A combination of glycerol and manganese promotes biofilm formation in Bacillus subtilis via the histidine kinase KinD signaling. J. Bacteriol.

Shen, A., and Higgins, D.E. (2006). The MogR transcriptional repressor regulates nonhierarchal expression of flagellar motility genes and virulence in Listeria monocytogenes. PLoS Pathog. *2*, e30.

Shen, A., Higgins, D.E., and Panne, D. (2009). Recognition of AT-rich DNA binding sites by the MogR repressor. Structure *17*, 769–777.

Sheppard, A.E., Poehlein, A., Rosenstiel, P., Liesegang, H., and Schulenburg, H. (2013). Complete Genome Sequence of Bacillus thuringiensis Strain 407 Cry-. Genome Announc. *1*.

Sierro, N., Makita, Y., de Hoon, M., and Nakai, K. (2008). DBTBS: a database of transcriptional regulation in Bacillus subtilis containing upstream intergenic conservation information. Nucleic Acids Res. *36*, D93–D96.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., *et al.* (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. *7*, 539.

Slamti, L., and Lereclus, D. (2002). A cell-cell signaling peptide activates the PlcR virulence regulon in bacteria of the Bacillus cereus group. EMBO J. *21*, 4550–4559.

Slamti, L., Perchat, S., Huillet, E., and Lereclus, D. (2014). Quorum sensing in Bacillus thuringiensis is required for completion of a full infectious cycle in the insect. Toxins (Basel). *6*, 2239–2255.

Slamti, L., Lemy, C., Henry, C., Guillot, A., Huillet, E., and Lereclus, D. (2016). CodY Regulates the Activity of the Virulence Quorum Sensor PlcR by Controlling the Import of the Signaling Peptide PapR in Bacillus thuringiensis. Front. Microbiol. *6*, 01501.

Smits, W.K., Bongiorni, C., Veening, J.-W., Hamoen, L.W., Kuipers, O.P., and Perego, M. (2007). Temporal separation of distinct differentiation pathways by a dual specificity Rap-Phr system in Bacillus subtilis. Mol. Microbiol. *65*, 103–120.

Söding, J. (2005). Protein homology detection by HMM-HMM comparison. Bioinformatics *21*, 951–960.

Solcan, N., Kwok, J., Fowler, P.W., Cameron, A.D., Drew, D., Iwata, S., and Newstead, S. (2012). Alternating access mechanism in the POT family of oligopeptide transporters. EMBO

J. *31*, 3411–3421.

Sonenshein, A.L. (2005). CodY, a global regulator of stationary phase and virulence in Gram-positive bacteria. Curr. Opin. Microbiol. *8*, 203–207.

Soutourina, O.A., Monot, M., Boudry, P., Saujet, L., Pichon, C., Sismeiro, O., Semenova, E., Severinov, K., Le Bouguenec, C., Coppée, J.-Y., *et al.* (2013). Genome-Wide Identification of Regulatory RNAs in the Human Pathogen Clostridium difficile. PLoS Genet. *9*, e1003493.

Souza, B.M., Castro, T.L. de P., Carvalho, R.D. de O., Seyffert, N., Silva, A., Miyoshi, A., and Azevedo, V. (2014). σ(ECF) factors of gram-positive bacteria: a focus on Bacillus subtilis and the CMNR group. Virulence *5*, 587–600.

Stachelhaus, T., Hüser, A., and Marahiel, M.A. (1996). Biochemical characterization of peptidyl carrier protein (PCP), the thiolation domain of multifunctional peptide synthetases. Chem. Biol. *3*, 913–921.

Stöver, A.G., and Driks, A. (1999a). Control of synthesis and secretion of the Bacillus subtilis protein YqxM. J. Bacteriol. *181*, 7065–7069.

Stöver, A.G., and Driks, A. (1999b). Regulation of synthesis of the Bacillus subtilis transition-phase, spore-associated antibacterial protein TasA. J. Bacteriol. *181*, 5476–5481.

Stragier, P., Kunkel, B., Kroos, L., and Losick, R. (1989). Chromosomal rearrangement generating a composite gene for a developmental transcription factor. Science *243*, 507–512.

Strauch, M.A. (1995a). In vitro binding affinity of the Bacillus subtilis AbrB protein to six different DNA target regions. J. Bacteriol. *177*, 4532–4536.

Strauch, M.A. (1995b). Delineation of AbrB-binding sites on the Bacillus subtilis spo0H, kinB, ftsAZ, and pbpE promoters and use of a derived homology to identify a previously unsuspected binding site in the bsuB1 methylase promote. J. Bacteriol. *177*, 6999–7002.

Strauch, M., Webb, V., Spiegelman, G., and Hoch, J.A. (1990). The SpoOA protein of Bacillus subtilis is a repressor of the abrB gene. Proc. Natl. Acad. Sci. U. S. A. *87*, 1801–1805.

Strauch, M.A., Wu, J.J., Jonas, R.H., and Hoch, J.A. (1993). A positive feedback loop controls transcription of the spoOF gene, a component of the sporulation phosphorelay in Bacillus subtilis. Mol. Microbiol. *7*, 967–974.

Stülke, J., Martin-Verstraete, I., Zagorec, M., Rose, M., Klier, A., and Rapoport, G. (1997). Induction of the Bacillus subtilis ptsGHI operon by glucose is controlled by a novel antiterminator, GlcT. Mol. Microbiol. *25*, 65–78.

Tatusov, R.L., Koonin, E. V, and Lipman, D.J. (1997). A genomic perspective on protein families. Science *278*, 631–637.

Tjalsma, H., Bolhuis, A., van Roosmalen, M.L., Wiegert, T., Schumann, W., Broekhuizen,

C.P., Quax, W.J., Venema, G., Bron, S., and van Dijl, J.M. (1998). Functional analysis of the secretory precursor processing machinery of Bacillus subtilis: identification of a eubacterial homolog of archaeal and eukaryotic signal peptidases. Genes Dev. *12*, 2318–2331.

Tojo, S., Hirooka, K., and Fujita, Y. (2013). Expression of kinA and kinB of Bacillus subtilis, necessary for sporulation initiation, is under positive stringent transcription control. J. Bacteriol. *195*, 1656–1665.

Tourasse, N.J., and Kolstø, A.-B. (2008). SuperCAT: a supertree database for combined and integrative multilocus sequence typing analysis of the Bacillus cereus group of bacteria (including B. cereus, B. anthracis and B. thuringiensis). Nucleic Acids Res. *36*, D461–D468.

Töwe, S., Leelakriangsak, M., Kobayashi, K., Van Duy, N., Hecker, M., Zuber, P., and Antelmann, H. (2007). The MarR-type repressor MhqR (YkvE) regulates multiple dioxygenases/glyoxalases and an azoreductase which confer resistance to 2-methylhydroquinone and catechol in Bacillus subtilis. Mol. Microbiol. *66*, 40–54.

Trach, K.A., and Hoch, J.A. (1993). Multisensory activation of the phosphorelay initiating sporulation in Bacillus subtilis: identification and sequence of the protein kinase of the alternate pathway. Mol. Microbiol. *8*, 69–79.

Trach, K., Burbulys, D., Strauch, M., Wu, J.-J., Dhillon, N., Jonas, R., Hanstein, C., Kallio, P., Perego, M., Bird, T., *et al.* (1991). Control of the initiation of sporulation in Bacillus subtilis by a phosphorelay. Res. Microbiol. *142*, 815–823.

U.S. Environmental Protection Agency (2010). Cry1Ab and Cry1F Bacillus thuringiensis (Bt) Corn Plant-Incorporated Protectants.

Underwood, S., Guan, S., Vijayasubhash, V., Baines, S.D., Graham, L., Lewis, R.J., Wilcox, M.H., and Stephenson, K. (2009). Characterization of the Sporulation Initiation Pathway of Clostridium difficile and Its Role in Toxin Production. J. Bacteriol. *191*, 7296–7305.

Vachon, V., Laprade, R., and Schwartz, J.-L. (2012). Current models of the mode of action of Bacillus thuringiensis insecticidal crystal proteins: a critical review. J. Invertebr. Pathol. *111*, 1–12.

Vanittanakom, N., Loeffler, W., Koch, U., and Jung, G. (1986). Fengycin--a novel antifungal lipopeptide antibiotic produced by Bacillus subtilis F-29-3. J. Antibiot. (Tokyo). *39*, 888–901.

Verhamme, D.T., Murray, E.J., and Stanley-Wall, N.R. (2009). DegU and Spo0A jointly control transcription of two loci required for complex colony development by Bacillus subtilis. J. Bacteriol. *191*, 100–108.

Verma, S.C., and Miyashiro, T. (2013). Quorum sensing in the squid-Vibrio symbiosis. Int. J. Mol. Sci. *14*, 16386–16401.

Verplaetse, E., Slamti, L., Gohar, M., and Lereclus, D. (2015). Cell Differentiation in a Bacillus thuringiensis Population during Planktonic Growth, Biofilm Formation, and Host

Infection. MBio *6*, e00138–15.

Vlamakis, H., Chai, Y., Beauregard, P., Losick, R., and Kolter, R. (2013). Sticking together: building a biofilm the Bacillus subtilis way. Nat. Rev. Microbiol. *11*, 157–168.

Walton, R.B., and Woodruff, H.B. (1949). A crystalline antifungal agent, mycosubtilin, isolated from subtilin broth. J. Clin. Invest. *28*, 924–926.

Wang, L., Grau, R., Perego, M., and Hoch, J.A. (1997). A novel histidine kinase inhibitor regulating development in Bacillus subtilis. Genes Dev. *11*, 2569–2579.

Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H.U., Bruccoleri, R., Lee, S.Y., Fischbach, M.A., Muller, R., Wohlleben, W., *et al.* (2015). antiSMASH 3.0--a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res. gkv437 – .

Weller, C., and Wu, M. (2015). A Generation Time Effect on the Rate of Molecular Evolution in Bacteria. Evolution *69*, 643–652.

York, K., Kenney, T.J., Satola, S., Moran, C.P., Poth, H., and Youngman, P. (1992). Spo0A controls the sigma A-dependent activation of Bacillus subtilis sporulation-specific transcription unit spoIIE. J. Bacteriol. *174*, 2648–2658.

Zhao, H., Msadek, T., Zapf, J., Madhusudan, Hoch, J.A., and Varughese, K.I. (2002). DNA complexed structure of the key transcription factor initiating development in sporulating bacteria. Structure *10*, 1041–1050.

Zouhir, S., Perchat, S., Nicaise, M., Perez, J., Guimaraes, B., Lereclus, D., and Nessler, S. (2013). Peptide-binding dependent conformational changes regulate the transcriptional activity of the quorum-sensor NprR. Nucleic Acids Res. *41*, 7920–7933.

ANNEXES

Erratum sur l’article : « Necrotrophism is a quorum-sensing-regulated lifestyle in *Bacillus thuringiensis*. »

Soumis à PLoS Pathogens

Dear Editor,

We would like to publish a correction about the following article :

Dubois, T., Faegri, K., Perchat, S., Lemy, C., Buisson, C., Nielsen-LeRoux, C., Gohar, M., Jacques, P., Ramarao, N., Kolsto, A-B.*, and Lereclus, D.* 2012. Necrotrophism is a quorum-sensing-regulated lifestyle in *Bacillus thuringiensis*. *PLoS Pathog.* **8:** e1002629. *co-corresponding authors.

In this paper we demonstrated that the quorum sensing regulator NprR was responsible for the capacity of *Bacillus thuringiensis* (*Bt*) to survive in the cadaver of infected insect larvae. In other terms, NprR was responsible for the necrotrophic lifestyle of the bacteria.

We reported experiments showing that a lipopeptide (called kurstakin) synthesized by a non-ribosomal peptide synthesis (NRPS) system was essential for the survival of the bacterial cells during their necrotrophic development. The NRPS system is encoded by a four-gene operon (*krsEABC*) whose expression is positively controlled by the quorum sensing regulator NprR. The role of the lipopeptide kurstakin in necrotrophism was shown by monitoring the total population of the bacterial cells in the insect cadavers and by comparing a *krsABC* mutant strain to the wild type *Bt* strain 407 Cry$^-$ (Figure 5B and 5C of the article). Similarly, the role of kurstakin in swarming mobility and biofilm formation was analyzed by comparing a *krsABC* mutant strain to the wild type *Bt* strain 407 Cry$^-$ (Figures 6A and 6B of the article). These errors are indicated in the pdf of the article.

It appears today that results described in Figures 5B, 5C (lower panel), 6A and 6B are not correct. Indeed, we have recently observed that the experiments on the role of the kurstakin resulted from a mistake. An analysis of the *krsABC* mutant strain revealed that the disruption of the *krsABC* by a *spc* gene, conferring resistance to spectinomycin, was introduced in the *Bacillus cereus* strain ATCC14579 instead of the *Bt* strain 407 Cry$^-$. The *B. cereus* strain ATCC14579 is closely related to the *Bt* strain 407 Cry$^-$ and is currently used in our laboratories for experiments focusing on *B. cereus*. A major difference between these two strains is the disruption of the *nprR* gene by a transposon in the *B. cereus* strain ATCC14579. As a result this strain has naturally a NprR$^-$ phenotype and should therefore be unable to survive in the insect cadavers like the *Bt* mutant strain 407 Δ*nprR* Δ*nprX* (see the article Dubois *et al.*, 2012, *PLoS Pathog.* **8:** e1002629: Fig. 1A and Materials and Methods section for the description of the *Bt* 407 Δ*nprR* Δ*nprX* mutant strain). Moreover, it was previously shown that the *B. cereus* strain ATCC14579 did not produce biofilm in LB peptone medium (Auger S, Krin E, Aymerich S, Gohar M. 2006. Autoinducer 2 affects biofilm formation by *Bacillus cereus*. *Appl Environ Microbiol*. 72:937-941) and did not swarm on EPS agar, a low-nutrient medium (Hsueh YH, Somers EB, Lereclus D, Ghelardi E, Wong AC. 2007. Biosurfactant production and surface translocation are regulated by PlcR in *Bacillus cereus* ATCC 14579 under low-nutrient conditions. *Appl Environ Microbiol*. 73:7225-7231).

The error we have done in experiments described in Figure 5B might therefore be due to the lack of a functional NprR regulator in the *B. cereus* strain ATCC14579 instead of the disruption of the *krsABC* genes in the *Bt* strain 407 Cry$^-$. In order to check this hypothesis, we attempted to disrupt the *krsABC* genes in the *Bt* strain 407 Cry$^-$. Several unsuccessful experiments suggested that it was impossible to disrupt these three genes by using our methodology. Thus, we decided to disrupt the *krsC* gene coding for a major synthetase of the NPRS system. The deletion of this gene abolishes the production of kurstakin (result not shown). To investigate the role of the kurstakin during the late stage of infection, we monitored the growth of the wild-type and Δ*krsC Bt* 407 Cry$^-$ strains in insect larvae (Figure 1). The results showed that the deletion of the *krsC* gene and the subsequent lack of kurstakin did not affect the survival of the *krsC* mutant strain for at least 3 days in the insect cadaver.

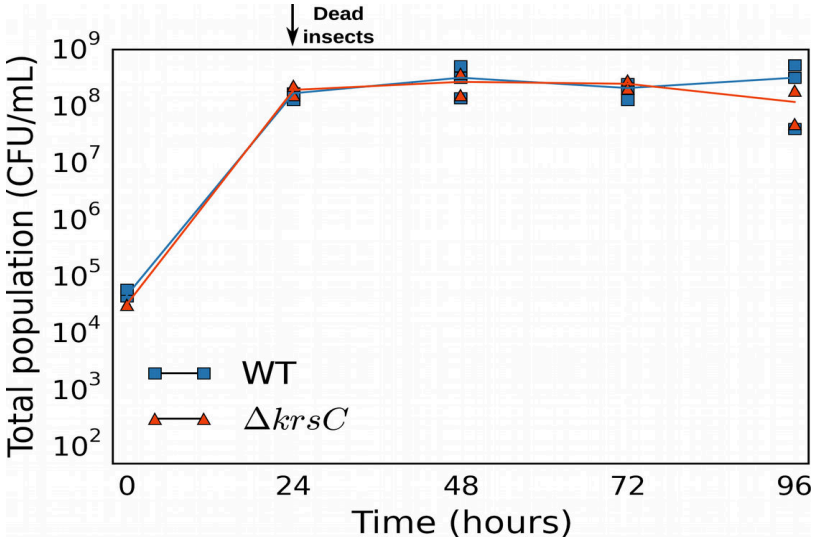


Figure 1. Kurstakin is not necessary for the survival of the *Bt* cells in insect cadavers.
Total bacterial CFU numeration after injection of *Bt* vegetative cells into the hemocoel of *Galleria mellonella* larvae. The population of the *Bt* 407 WT and Δ*krsC mutant* strains are maintained at the same level for at least 3 days after the death of the insect larvae.

These new data show that the lipopeptide kurstakin is not necessary for the survival of the wild type *Bt* strain 407 Cry⁻ in the insect cadavers. However, we have repeated the experiment showing that insertion of a constitutive promoter upstream of the *krs* operon in the Δ*nprR* Δ*nprX* mutant strain (ΔRX P_{aphA3}-*krs* strain) partially and significantly restored the survival of the bacteria in the insect cadaver (Figure 2). The results suggest that kurstakin, although not necessary for the survival of the bacteria, plays an important role in this property of the *Bt* cells.

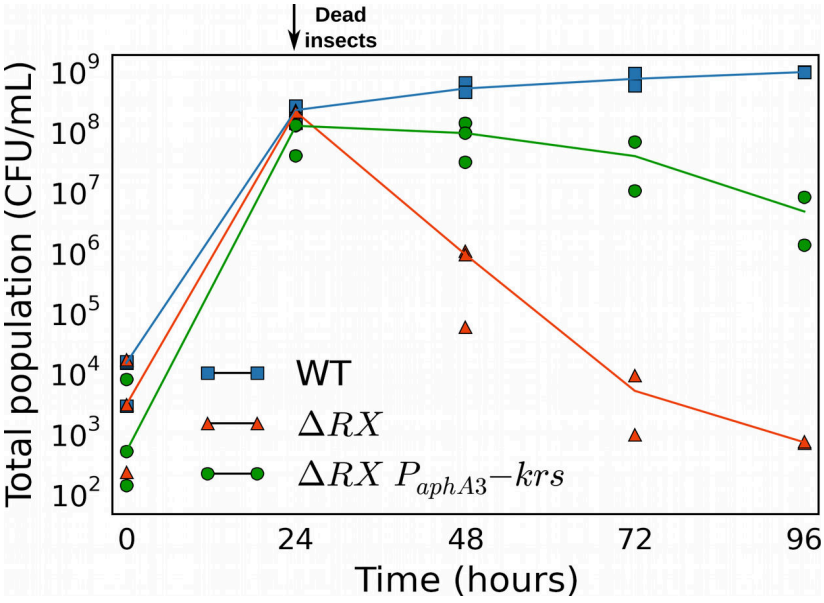


Figure 2. Kurstakin plays a significant role in the necrotrophic lifestyle of *Bt*.
Total bacterial CFU numeration after injection of *Bt* vegetative cells into the hemocoel of *Galleria mellonella* larvae. The constitutive expression of the *krsABC* genes in the ΔRX strain partially restored the survival of this mutant.

These new results were obtained and submitted by:

Sébastien Gélis-Jeanvoine[1,2], Christelle Lemy[1,2], Thibault Caradec[3], Leyla Slamti[1,2], Michel Gohar[1,2], Philippe Jacques[3] Anne-Brit Kolstø[4*] and Didier Lereclus[1,2*]

[1]INRA, UMR1319 Micalis, F-78352 Jouy-en-Josas, France

[2]AgroparisTech, UMR Micalis, F-78352 Jouy-en-Josas, France

[3]ProBioGEM, UPRES EA 1026, Université Lille Nord de France, USTL, F-59655 Villeneuve d'Ascq, France

[4]School of Pharmacy, Laboratory for Microbial Dynamics, University of Oslo, N-0316 Oslo, Norway

[*]Corresponding authors

Please let us know if this letter is sufficient or if we should provide these new data in a form more appropriate to the guidelines of PloS pathogens.

We hope that this correction can be published on line and linked to the article *PLoS Pathog.* **8:** e1002629 (Dubois *et al.*, 2012).

Sincerely yours,

Didier Lereclus and Anne-Brit Kolstø

Title: A study about kurstakin and Spo0A, two adaptation factors of *Bacillus thuringiensis*.

Keywords: *Bacillus thuringiensis*, Transcriptional regulation, Necrotrophism, Spo0A, Kurstakin, Biofilm.

Abstract:

Bacillus thuringiensis is an entomopathogenic bacterium from the *Bacillus cereus* group. Additionally to its ability to kill larvae of susceptible insects, *B. thuringiensis* is able to maintain a vegetative population in its dead host for more than a week; a behavior called necrotrophism. Like other *Bacillus*, *B. thuringiensis* is also capable of triggering several adaptation processes: biofilm formation, sporulation and synthesis of original secondary metabolites for instance. These processes are tightly regulated by environmental cues sensed by several sensors of extra-cellular signals, often linked to master regulators of transcription, among which are Spo0A and NprR. Spo0A is notably present in the *Bacillus* and *Clostridium* species and controls numerous adaptation processes. NprR is a *quorum sensing* regulator specific to the *Bacillus cereus* group. In *B. thuringiensis*, when the *quorum* is reached, it allows the transition between a pathogenic and a necrotrophic lifestyle. The work undertaken during this thesis aimed at a better understanding of the behavior of *B. thuringiensis* during the stationary phase and the infection of a host insect.

NprR regulates the transcription of about forty genes, among which are those of the *krs* locus. This big 30 kb locus harbors six genes involved in the non-ribosomal synthesis of kurstakin, a lipo-heptapeptide possibly playing an important role in necrotrophism, and exhibiting biotechnological potential. We showed that the genes are co-transcribed from a promoter upstream of the first gene, forming an operon. Measures of this promoter's activity revealed a strong dependence to growth conditions and a complex regulation involving, in addition to NprR, Spo0A and a DNA palindrome of yet unknown function. We also show that the product of this locus, kurstakin, is involved in the structuring of the biofilm of *B. thuringiensis*.

Using a broader approach, we characterized the *in vitro* (two time points) and *in vivo* transcriptomes of our model strain 407 and its *spo0A*-deletion mutant, using RNA-seq. Coupled to an *in silico* prediction of the Spo0A boxes in the genome of *B. thuringiensis*, these data describe for the first time the Spo0A regulon of a member of the *Bacillus cereus* group. We also investigate the role of Spo0A in the infection cycle of *B. thuringiensis* and show that a *spo0A*-less mutant of *B. thuringiensis* is unable to persist in its host cadaver. This survival defect does not depend on sporulation, because a mutant impaired in later steps of sporulation still has the ability to survive. This suggests that part of the Spo0A regulon is also involved in the necrotrophic behavior of *B. thuringiensis*.

Titre : Étude de la kurstakine et de Spo0A, deux facteurs d'adaptation de *Bacillus thuringiensis*.

Mots clés : *Bacillus thuringiensis*, Régulation transcriptionnelle, Nécrotrophisme, Spo0A, Kurstakine, Biofilm.

Résumé :

Bacillus thuringiensis est une bactérie entomopathogène du groupe *Bacillus cereus*. En plus de sa capacité à infecter et tuer les larves d'insectes sensibles, *B. thuringiensis* peut survivre à l'état végétatif plus d'une semaine dans le cadavre de l'hôte qu'il vient de tuer ; un mode de développement appelé nécrotrophisme. Comme d'autres *Bacillus*, *B. thuringiensis* possède aussi plusieurs mécanismes d'adaptation : formation de biofilm, sporulation et synthèse de métabolites secondaires originaux, par exemple. Ceux-ci sont finement régulés en fonction de l'environnement, *via* des systèmes d'intégration de signaux extra-cellulaires qui convergent souvent vers des régulateurs transcriptionnels centraux, parmi lesquels Spo0A et NprR. Spo0A est notamment présent chez les espèces *Bacillus* et *Clostridium*, et contrôle un grand nombre de mécanismes d'adaptation. NprR est un régulateur de *quorum sensing* présent uniquement dans le groupe *Bacillus cereus*. Chez *B. thuringiensis*, lorsque le *quorum* est atteint, il permet la transition d'un mode de vie pathogène à nécrotrophe. Le travail mené au cours de cette thèse a eu pour but de mieux comprendre le comportement de *B. thuringiensis* en phase stationnaire et au cours de l'infection d'un insecte hôte.

NprR régule la transcription d'une quarantaine de gènes, parmi lesquels le locus *krs*. Ce grand locus de 30 kb porte six gènes impliqués dans la synthèse non-ribosomique de la kurstakine, un lipo-heptapeptide susceptible de jouer un ou des rôle(s) importants dans le nécrotrophisme et revêtant un potentiel intérêt bio-technologique. Nous avons montré que les gènes sont co-transcrits à partir d'un promoteur en amont du premier gène, formant donc un opéron. La mesure de l'activité de ce promoteur nous a permis de mettre en évidence que celle-ci est fortement dépendante des conditions de croissance, et que sa régulation est complexe et fait intervenir, en plus de NprR, Spo0A et une séquence d'ADN palindromique dont la fonction est encore inconnue. Nous démontrons aussi que le produit de ce locus, la kurstakine, est impliqué dans la structuration du biofilm de *B. thuringiensis*.

Dans le cadre d'une approche plus globale, nous avons caractérisé les transcriptomes de *B. thuringiensis in vitro*, à deux points dans le temps, et *in vivo*, dans la souche modèle 407 et le mutant de délétion de *spo0A*, en utilisant la technique de RNA-seq. Couplées à une prédiction *in silico* des boîtes Spo0A dans le génome de *B. thuringiensis*, ces données décrivent pour la première fois le régulon Spo0A d'une bactérie du groupe *Bacillus cereus*. Nous tentons également d'élucider le rôle de Spo0A dans le cycle infectieux de *B. thuringiensis*, et montrons que sans *spo0A*, *B. thuringiensis* est incapable de persister dans le cadavre de son hôte. Ce défaut de survie est indépendant de la sporulation, car un mutant affecté dans les étapes plus tardives de la sporulation survit. Cela suggère qu'une partie du régulon Spo0A est également impliquée dans mode de vie nécrotrophe de *B. thuringiensis*.