



**HAL**  
open science

# Mécanismes moléculaires de la survie à long terme chez *Propionibacterium freudenreichii*

Flavia Figueira Aburjaile

► **To cite this version:**

Flavia Figueira Aburjaile. Mécanismes moléculaires de la survie à long terme chez *Propionibacterium freudenreichii*. Alimentation et Nutrition. Agrocampus Ouest, 2015. Français. NNT : 2015NSARB273 . tel-01697901v2

**HAL Id: tel-01697901**

**<https://theses.hal.science/tel-01697901v2>**

Submitted on 1 Feb 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## RÉSUMÉ ABSTRACT

Mécanismes moléculaires de la survie à long terme chez *Propionibacterium freudenreichii*

*Propionibacterium freudenreichii* est une bactérie très utilisée par l'industrie laitière. Elle appartient aux Actinomycètes connus pour leur survie pendant de longues périodes, dans des conditions environnementales défavorables. Pour mieux comprendre ce phénomène, la caractérisation phénotypique de 8 souches de *P. freudenreichii* a été réalisée sur 11 jours dans un milieu en carence nutritionnelle. Le taux de survie bactérienne a été mesuré par densité optique, par énumération et évaluation de la viabilité cellulaire. En outre, l'absence de lyse cellulaire a été évaluée par PCR quantitative. La croissance de *P. freudenreichii* a été décrite en phases exponentielle, stationnaire, stationnaire tardive et survie à long terme. Dans nos conditions expérimentales pendant la période de survie à long terme, les bactéries sont restées viables. La caractérisation phénotypique a montré que *P. freudenreichii* CIRM-BIA138 était la plus résistante à la carence nutritionnelle et entrainait dans un état viable mais non-cultivable. Cette souche a été utilisée pour une étude fonctionnelle par RNA-Seq ainsi que pour des analyses biochimiques sur les surnageants de culture, en phases exponentielle et stationnaire. L'association de ces données transcriptomiques et métabolomiques a permis de déduire les stratégies impliquées dans la survie de cette bactérie. La préparation à l'état de dormance, la diminution du métabolisme et l'utilisation de sources alternatives d'énergie semblent impliquées dans l'adaptation et la persistance de *P. freudenreichii* CIRM-BIA138 en carence nutritionnelle durant de longues périodes.

**Mots-clés :** *P. freudenreichii*, survie à long terme, RNA-Seq, métabolome, VBNC.

*Molecular mechanisms of long-term survival in Propionibacterium freudenreichii*

*Propionibacterium freudenreichii* is a dairy bacterium belonging to the Actinobacteria group, which is known to survive for long periods in harsh environmental conditions. In order to investigate the long-term survival phenomenon in *P. freudenreichii*, 8 strains were phenotypically characterized for a period of 11 days in nutrient shortage condition. Bacterial survival rate was assessed by optical density, CFU counting and live-dead cellular viability. In addition, the absence of cell lysis was evaluated by quantitative PCR. *P. freudenreichii* growth phases were classified as exponential, stationary, late stationary and long-term survival. Moreover, it was observed that bacterial viability was maintained during long-term survival. Phenotypical characterization indicated that *P. freudenreichii* CIRM-BIA138 was more resistant to nutrient shortage being able to enter into a viable but nonculturable dormant state. In addition, functional studies of this strain were conducted by RNA-Seq on cultures sampled in exponential and stationary growth phases. Concomitantly, several biochemical analyses were carried out on the culture supernatant. An integrative approach of metabolomic and transcriptomic data allowed us to infer strategies associated with the survival of this bacterium, such as preparation for the dormant state, slow down of metabolic activity and utilization of alternative sources of energy, which altogether might allow *P. freudenreichii* CIRM-BIA 138 to adapt and persist through nutrient shortage for long periods.

**Keywords:** *Propionibacterium freudenreichii*, long-term survival, RNA-Seq, metabolomics, viable but nonculturable.

Mécanismes moléculaires de la survie à long terme chez *Propionibacterium freudenreichii*

# AGRO CAMPUS

## OUEST



UNIVERSITÉ  
EUROPÉENNE  
DE BRETAGNE

**Flavia FIGUEIRA ABURJAILE • 9 décembre 2015**

Thèse AGROCAMPUS OUEST  
sous le label de l'Université Européenne de Bretagne pour  
obtenir le titre de  
**DOCTEUR D'AGROCAMPUS OUEST**  
Spécialité Biochimie, Biologie moléculaire et cellulaire

ÉCOLE DOCTORALE • Vie-Agro-Santé (VAS)

LABORATOIRE D'ACCUEIL • UMR 1253 INRA - AGROCAMPUS  
OUEST Science et Technologie du Lait et de l'Œuf (STLO)  
THÈSE EN COTUTELLE • UFMG, Belo Horizonte, Brésil



AGROCAMPUS OUEST • Institut supérieur des sciences agronomiques,  
agroalimentaires, horticoles et du paysage  
65 rue de Saint-Brieuc - CS84215 - F-35042 Rennes Cedex  
Tél. : 02 23 48 50 00  
www.agrocampus-ouest.fr



Thèse B-273 — 2015-28 • FIGUEIRA Flavia



## Mécanismes moléculaires de la survie à long terme chez *Propionibacterium freudenreichii*

**Michel GAUTIER**  
Professeur, AGROCAMPUS OUEST, UMR INRA AO STLO /  
président

**Muriel COCAIGN-BOUSQUET**  
Directrice de recherche, INRA Toulouse / rapporteur

**Artur SILVA**  
Professeur, université fédérale de Parà, Brésil / rapporteur

**Marie-France PILET**  
Maître de conférences, ONIRIS, Nantes / examinatrice

**Vasco Ariston de CARVALHO AZEVEDO**  
Professeur, UFMG Belo Horizonte, Brésil / directeur de thèse

**Yves LE LOIR**  
Directeur de recherche, UMR INRA-AO STLO / directeur de thèse

**Hélène FALENTIN**  
Ingénieure de recherche, UMR INRA-AO STLO / co encadrante



N° ordre : 2015-28

N° Série : B-273

## **THESE / AGROCAMPUS OUEST**

Sous le label de l'Université Européenne de Bretagne

pour obtenir le diplôme de :

**DOCTEUR DE L'INSTITUT SUPERIEUR DES SCIENCES AGRONOMIQUES,  
AGRO-ALIMENTAIRES, HORTICOLES ET DU PAYSAGE**

Spécialité : Spécialité Biochimie, Biologie moléculaire et cellulaire

**Ecole Doctorale : « Vie Agro Santé »**

présentée par :

**Flavia Figueira Aburjaile**

**Mécanismes moléculaires de la survie à long terme chez**

***Propionibacterium freudenreichii***

soutenue le 9 décembre 2015 devant la commission d'Examen

Composition du jury :

Rapporteurs :

Muriel Coccagn-Bousquet (Directrice de recherche, INRA, Toulouse, France)

Artur Silva (Professeur, UFPA, Belém, Brésil)

Membres :

Marie-France Pilet (Maître de conférences, ONIRIS, Nantes; Examinatrice)

Michel Gautier (Professeur, AGROCAMPUS OUEST, Rennes, France; Président)

Directeurs de Thèse :

Vasco Ariston de Carvalho Azevedo (Professeur, UFMG, Belo Horizonte, Brésil)

Yves Le Loir (Directeur de recherche, INRA, Rennes, France)

Co-encadrante de Thèse :

Hélène Falentin (Ingénieur de recherche, INRA, Rennes, France)



UNIVERSIDADE FEDERAL DE MINAS GERAIS  
INSTITUTO DE CIÊNCIAS BIOLÓGICAS  
PROGRAMA INTERUNIDADES DE PÓS-GRADUAÇÃO  
EM BIOINFORMÁTICA

**Tese**

**Mecanismos moleculares de sobrevivência em longo prazo em  
*Propionibacterium freudenreichii***

**Orientada:** Flavia Figueira Aburjaile

**Orientadores:** Prof. Dr. Vasco Ariston de Carvalho Azevedo

Dr. Yves Le Loir

**Rennes**

**2015**



Flavia Figueira Aburjaile



Mecanismos moleculares de sobrevivência em longo prazo em  
*Propionibacterium freudenreichii*

Tese apresentada ao Programa Interunidades de Pós-Graduação em Bioinformática da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título de Doutor em Bioinformática.

**Orientadores:** Prof. Dr. Vasco Azevedo  
Dr. Yves Le Loir

**Rennes**  
**2015**

*“Foi o tempo que dedicaste à tua rosa que a fez tão importante”.*  
*« C'est le temps que tu as perdu pour ta rose qui fait ta rose si importante ».*

**Antoine de Saint-Exupéry**

## **AGRADECIMENTOS**

- *Agradeço a Deus por ter me dado persistência e determinação durante estes quatro anos de trabalho seguindo em frente com os meus objetivos e não desanimando com as dificuldades;*
- *Ao meu orientador brasileiro, Vasco Azevedo, por todo o aprendizado e ensinamento durante estes anos;*
- *Ao meu orientador francês, Yves Le Loir, pelo recebimento em seu Instituto de Pesquisa;*
- *Aos membros da banca pela disponibilidade em avaliarem este trabalho de Tese;*
- *A H  l  ne Falentin e aos membros do INRA de Rennes, pela ajuda e disponibilidade;*
- *A Anne Pinto por todo o acolhimento, paci  ncia e ajuda durante a reda  o da Tese;*
- *Paulette e Jessica, t  cnicas, pela prepara  o de todos os meios de cultura e de materiais essenciais utilizados nos experimentos;*
- *   P  s-gradua  o em Bioinform  tica, por todo o suporte na parte administrativa do processo desde o in  cio da sele  o de doutorado;*
- *As ag  ncias de fomento CNPq, CAPES e CAPES-COFECUB pelo aux  lio financeiro concedido no Brasil e na Fran  a durante este per  odo de Tese;*
- *A todos os “LGCMistas”, pelos aprendizados dentro e fora do laborat  rio, pelas amizades e pelas conquistas. N  o vou descrever nomes, pois voc  s todos s  o muito especiais!*
- *Aos meus colegas e amigos que fiz durante a minha caminhada em Rennes, estrangeiros e brasileiros, aos la  os que se criaram e que continuar  o, com certeza, nos pr  ximos anos. E a frase que resume toda esta trajet  ria, est   em meu livro favorito, O pequeno pr  ncipe: “Tu te tornas eternamente respons  vel por aquilo que cativas (Antoine de Saint-Exup  ry)”;*
- *Aos meus queridos amigos e amigas de inf  ncia, de col  gio, de faculdade e de est  gio, que me apoiaram em todos os momentos, de perto ou de longe*

*- E principalmente à minha família, pelo apoio e amor incondicional em todos os momentos, em especial aos meus pais, Regina e Samir e aos “meus gêmeos”, queridos irmãos: Arthur e Renata.*



## REMERCIEMENTS

*Je remercie :*

- *Dieu de m'avoir donné la persistance et la détermination au cours de ces quatre années de travail, qui m'ont permis de suivre sans cesse mes objectifs et de ne pas baisser les bras face aux difficultés;*
- *mon directeur de Thèse brésilien, Vasco Azevedo, tout l'apprentissage et l'enseignement au long de ces années;*
- *mon directeur de Thèse français, Yves Le Loir, pour l'accueil au sein de l'Institut de recherche;*
- *les membres du jury, leur disponibilité dans l'évaluation de cette Thèse;*
- *Hélène Falentin et aux membres d'INRA de Rennes, pour l'aide et la disponibilité;*
- *Anne Pinto pour l'accueil, la patience et l'aide pendant la rédaction de la Thèse;*
- *Paulette et Jessica, techniciens, pour la préparation de tous les milieux de culture et des matériaux essentiels utilisés dans les manipulations;*
- *le secrétariat de Bioinformatique, pour le support dans la partie administrative du processus depuis le début de la sélection de doctorat;*
- *les organismes de financement CNPq, CAPES et CAPES-COFECUB, le soutien financier accordé au Brésil et en France au cours de cette période de Thèse;*
- *tous les étudiants du laboratoire brésilien, les «LGCMistas», grâce à l'expérience acquise à l'intérieur et à l'extérieur du laboratoire, à l'amitié et aux réalisations. Je ne vais pas citer tous vos noms, car vous êtes tous très importants!*
- *les collègues et amis que j'ai connus lors de mon séjour à Rennes, étrangers et brésiliens, ces amitiés qui ont été créés se poursuivront, bien sûr, dans les prochaines années. Et la phrase qui peut résumer toute cette trajectoire, extraite de mon livre préféré, Le Petit Prince: « Tu deviens responsable pour toujours de ce que tu as apprivoisé » (Antoine de Saint-Exupéry);*
- *mes chers amis et amies d'enfance, de collège, de faculté et de stage, qui m'ont toujours soutenue, de près ou de loin*

- et surtout ma famille pour leur soutien et leur amour inconditionnel, mes parents surtout, Regina et Samir et « mes jumeaux », chers frère et soeur : Arthur et Renata.

## RESUMO

*Propionibacterium freudenreichii* é uma bactéria amplamente utilizada na indústria de laticínios, na síntese de biomoléculas e na fabricação de produtos probióticos. Esta bactéria é pertencente ao grupo de Actinobacterias, que são conhecidas por sua sobrevivência por períodos prolongados a condições ambientais adversas. Para uma melhor compreensão do fenômeno de sobrevivência à longo-prazo, a caracterização fenotípica de 8 linhagens de *P. freudenreichii* foi realizada neste trabalho durante um período de 11 dias em um meio de carência nutricional. A taxa de sobrevivência bacteriana foi monitorada através da densidade óptica, da enumeração de colônias e da viabilidade celular (kit live/dead). Além disso, a ausência de lise celular foi avaliada utilizando-se a *Real-time quantitative PCR (qPCR)*. Neste trabalho, as fases de crescimento de *P. freudenreichii* foram descritas e classificadas em: exponencial, estacionária, estacionária tardia e sobrevivência à longo-prazo. Em nossas condições experimentais, observou-se que durante o período de sobrevivência à longo prazo, as bactérias mantiveram-se viáveis. A caracterização fenotípica indica que, dentre as 8 linhagens, *P. freudenreichii* CIRM-BIA 138 demonstrou maior resistência a escassez nutricional e suas células entraram em um estado de dormência denominado viáveis, mas não cultiváveis (VBNC). Esta linhagem CIRM-BIA 138 foi utilizada para um estudo funcional através da abordagem de RNA-Seq, em fase de crescimento exponencial e estacionária. Concomitante, um conjunto de análises bioquímicas (dosagem de aminoácidos, de ácidos orgânicos e de açúcares) foi realizado a partir do sobrenadante durante a cinética bacteriana. A associação destes dados transcriptômicos e metabolômicos permitiu inferir estratégias envolvidas na sobrevivência desta bactéria. A preparação para um estado de dormência, a diminuição do metabolismo e a utilização de fontes alternativas de energia estão envolvidas na adaptação e na persistência de *P. freudenreichii* CIRM-BIA 138 em escassez nutricional durante longos períodos.

**Palavras-chave:** *Propionibacterium freudenreichii*, mecanismos de sobrevivência à longo prazo, RNA-Seq, metabolômica, VBNC.

## ABSTRACT

*Propionibacterium freudenreichii* is a bacterium widely used in the dairy industry, for biomolecule synthesis, and in the probiotic market. This species belongs to the Actinobacteria group, which is known to survive for long periods in harsh environmental conditions. Therefore, in order to better understand the long-term survival phenomenon of *P. freudenreichii*, 8 strains were phenotypically characterized in a laboratory medium for a period of 11 days, in condition of nutrient shortage. Bacterial survival rate was assessed by optical density, plate counting, and by assessing cellular viability (live/dead). In addition, the absence of cell lysis was evaluated by quantitative PCR (qPCR). The *P. freudenreichii* growth phases were classified as exponential, stationary, late stationary and long-term survival. Under the experimental conditions, it was observed that bacterial viability was maintained during the long-term survival. Phenotypical characterization indicated that *P. freudenreichii* CIRM-BIA138 was more resistant to nutrient shortage and that this strain was able to enter into a viable but nonculturable dormant state (VBNC). Therefore, functional studies of this strain were conducted by RNA-Sequencing on cultures sampled in the exponential and stationary growth phases. Concomitantly, a series of biochemical studies (quantitation of amino acids, organic acids and sugars) were carried out on the bacterial supernatant. The integration of metabolome and transcriptome data allowed us to infer strategies associated with the survival of this bacterium, such as preparation for the dormant state, slow down of the metabolic activity and utilization of alternative sources of energy, which altogether might allow *P. freudenreichii* CIRM-BIA 138 to adapt and persist throughout nutrient shortage for long periods.

**Keywords:** *Propionibacterium freudenreichii*, long-term survival mechanism, RNA-Seq, metabolomics, VBNC.

## RÉSUMÉ

La *Propionibacterium freudenreichii* est une bactérie largement utilisée dans l'industrie laitière, dans la synthèse de biomolécules et dans la fabrication de produits probiotiques. Cette bactérie appartient au groupe des Actinomycètes, qui sont connues pour leur survie pendant des périodes prolongées dans des conditions environnementales défavorables. Pour mieux comprendre ce phénomène de survie à long terme, la caractérisation phénotypique de 8 souches de *P. freudenreichii* a été réalisée dans ce travail sur une période de 11 jours dans un milieu en carence nutritionnelle. Le taux de survie bactérienne a été mesuré par densité optique, par énumération et évaluation de la viabilité cellulaire (live/dead). En outre, l'absence de lyse cellulaire a été évaluée par PCR quantitative (qPCR). Dans ce travail, la croissance de *P. freudenreichii* ont été a été décrite et divisée en phases exponentielle, stationnaire, stationnaire tardive et survie à long terme. Dans nos conditions expérimentales, il a été observé que, pendant la période de survie à long terme, les bactéries restaient viables. La caractérisation phénotypique des 8 souches a montré que *P. freudenreichii* CIRM-BIA138 était la plus résistante à la carence nutritionnelle et que les cellules bactériennes entraient dans un état de dormance qualifié d'état viable mais non-cultivable (VBNC). Cette souche CIRM-BIA138 a été utilisée pour une étude fonctionnelle par une approche RNA-Seq, sur les phases de croissance exponentielle et stationnaire. Parallèlement, un ensemble d'analyses biochimiques (dosage d'acides aminés, d'acides organiques et de sucres) a été réalisé sur les surnageants de culture au cours de la cinétique. L'association de ces données transcriptomiques et métabolomiques a permis de déduire quelques stratégies impliquées dans la survie de cette bactérie. La préparation à l'état de dormance, la diminution du métabolisme et l'utilisation de sources alternatives d'énergie semblent impliqués dans l'adaptation et la persistance de *P. freudenreichii* CIRM-BIA138 en carence nutritionnelle pendant de longues périodes.

**Mots-clés :** *Propionibacterium freudenreichii*, mécanismes de survie à long terme, RNA-Seq, métabolomique, VBNC.

## Sumário / Sommaire

LISTA DE FIGURAS .....	15
LISTE DES FIGURES .....	16
LISTA DE TABELAS .....	17
LISTE DES TABLEAUX.....	18
LISTA DE ABREVIATURAS .....	19
LISTE DE ABREVIATIONS UTILISEES .....	21
APRESENTAÇÃO / PRÉSENTATION .....	23
Colaborações .....	24
Collaborations.....	25
Prefácio .....	26
Préambule .....	29
Objetivos / Objectifs .....	32
Objetivos.....	33
Objectifs.....	34
CAPÍTULO 1: REVISÃO DA LITERATURA.....	35
CHAPITRE 1 : RÉVISION DE LA LITTÉRATURE.....	35
1. Gênero <i>Propionibacterium</i> .....	36
1.1. Ecologia, fisiologia e classificação .....	36
1.2. Caracterização e importância do microrganismo em estudo: <i>Propionibacterium freudenreichii</i> .....	37
2. Aplicações industriais de <i>Propionibacterium freudenreichii</i> .....	39
2.1. Indústria de queijos PPC (“pasta cozida e prensada”) .....	42
2.2. Otimização da produção de vitamina B12 .....	48
2.3. Obtenção do ácido propiônico.....	50
2.4. Probiótico .....	52
2.4.1. Histórico e definição do termo probiótico .....	52
2.4.2. Controle da microbiota .....	53
2.4.3. Imunomodulação.....	55
2.4.4. Apoptose induzida por ácidos graxos de cadeia curta .....	56
2.4.5. Molécula ácido 1,4-di-hidroxi-2-naftóico (DHNA) .....	58
2.4.6. Capacidade de adaptação .....	59
3. Genoma de <i>Propionibacterium freudenreichii</i> .....	61
4. Transcriptômica de <i>Propionibacterium freudenreichii</i> .....	65

5. Principais vias metabólicas de <i>Propionibacterium freudenreichii</i> .....	67
INTRODUCTION BIBLIOGRAPHIQUE.....	69
1. Genre <i>Propionibacterium</i> .....	70
1.1. Classification, écologie et physiologie.....	70
1.2. La caractérisation et l'importance du micro-organisme étudié : <i>Propionibacterium freudenreichii</i> .....	71
2. Applications industrielles de <i>Propionibacterium freudenreichii</i> .....	73
2.1. Industrie des fromages PPC (pâte pressée cuite) .....	76
2.2. Optimisation de la production de vitamine B12.....	82
2.3. L'obtention d'acide propionique.....	84
2.4. Probiotique .....	86
2.4.1 L'histoire et la définition du terme probiotique .....	86
2.4.2 Contrôle du microbiote .....	87
2.4.3. Immunomodulation .....	89
2.4.4. Apoptose induite par les acides gras à chaîne courte.....	90
2.4.5. Molécule acide 1,4-dihydroxy-2-naphtoïque (DHNA) .....	92
2.4.6. Capacité d'adaptation .....	93
3. Génomique de <i>Propionibacterium freudenreichii</i> .....	95
4. Transcriptomique de <i>Propionibacterium freudenreichii</i> .....	98
5. Les principales voies métaboliques de <i>Propionibacterium freudenreichii</i> .....	100
CAPÍTULO 2 : RESULTADOS .....	102
CHAPITRE 2 : RÉSULTATS .....	102
SEÇÃO I: Caracterização do fenótipo de linhagens de <i>Propionibacterium freudenreichii</i> .....	103
SECTION I : Caractérisation du phénotype des souches de <i>Propionibacterium</i> <i>freudenreichii</i> .....	103
1. Introdução .....	104
1. Introduction.....	106
2. Artigo 1: The long-term survival of <i>Propionibacterium freudenreichii</i> in a context of nutrient shortage. ....	108
SEÇÃO II: Utilização da genômica funcional para a compreensão do fenômeno de sobrevivência à longo prazo em <i>Propionibacterium freudenreichii</i> .....	134
SECTION II : L'utilisation de la génomique fonctionnelle pour comprendre le phénomène de la survie à long terme chez <i>Propionibacterium freudenreichii</i> .....	134
1. Introdução.....	135

1. Introduction .....	137
2. Artigo 2: Adaptation of <i>Propionibacterium freudenreichii</i> to long-term survival in case of gradual nutritional shortage .....	139
CAPÍTULO 3: Discussão geral, conclusão e perspectivas.....	258
CHAPITRE 3 : Discussion générale, conclusion et perspectives .....	258
Discussão geral .....	259
Discussion générale .....	264
Conclusão .....	271
Conclusion .....	272
Perspectivas .....	274
Perspectives .....	278
Referências/ Références .....	283
Anexos / Annexes .....	310
Anexo 1: Capítulo de livro: <i>Genomics</i> publicado em <i>A Text Book of Biotechnology</i> (disponível em <a href="http://www.smgebooks.com">www.smgebooks.com</a> ) .....	311
Annexe 1 : Chapitre du livre : <i>Genomics</i> publié dans <i>A Text Book of Biotechnology</i> (disponible en <a href="http://www.smgebooks.com">www.smgebooks.com</a> ) .....	311
Anexo 2: Capítulo de livro: <i>RNA-seq: Reveling biological insights in Bacteria</i> aceito em <i>Next Generation Sequencing</i> (disponível em breve em InTech - Open Access Publisher).....	333
Annexe 2 : Chapitre du livre : <i>RNA-seq: Reveling biological insights in Bacteria</i> accepté dans <i>Next Generation Sequencing</i> (bientôt disponible sur InTech - Open Access Publisher).....	333
Anexo 4: Nova montagem e anotação do genoma de <i>Propionibacterium freudenreichii</i> CIRM-BIA138.....	377
Annexe 4 : Nouvel assemblage et annotation du génome de <i>Propionibacterium freudenreichii</i> CIRM-BIA138.....	378
Anexo 5: Dados genômicos do <i>Genomes Online Database</i> (GOLD) .....	379
Annexe 5 : Données génomiques du <i>Genomes Online Database</i> (GOLD).....	381
Anexo 6: Manuscrito em redação - Survival strategies of Gram-positive bacteria ..	383
Annexe 6 : Manuscrit en cours - Survival strategies of Gram-positive bacteria ....	383
Anexo 7: Curriculum vitae.....	401
Annexe 7 : Curriculum vitae.....	401



## LISTA DE FIGURAS

Figura 1: Aspecto pleomórfico de <i>P. freudenreichii</i> .....	37
Figura 2: Produção de leite de vaca no território francês.....	42
Figura 3: Processos de transformação de queijos PPC.....	45
Figura 4: <i>P. freudenreichii</i> em um queijo Emmental.....	46
Figura 5: Esquema da reação de hidrólise de triacilglicérides.....	46
Figura 6: Representação da fórmula química do ácido propiônico.....	50
Figura 7: Publicações no PubMed com as palavras-chave “probiotics bacteria” ao decorrer dos últimos anos.....	53
Figura 8: Espécies microbianas dominantes em porções do trato gastrointestinal.....	54
Figura 9: Indução da apoptose a partir de SCFAs produzidos por bactérias propiônicas no cólon.....	56
Figura 10: Combinação dos dados genômicos, transcriptômicos e proteômicos.....	66
Figura 11: Via metabólica de Wood-Werkman em <i>P. freudenreichii</i> acoplada ao ciclo de Krebs.....	67
Figura 12: Esquema de processos metabólicos envolvidos na maturação dos queijos PPC visando a obtenção de compostos aromáticos.....	68
Figura 13: Número total de projetos por ano depositados no GOLD.....	379
Figura 14: Distribuição de projetos genômicos em bactérias de acordo com sua classificação.....	379
Figura 15: Relevância dos projetos bacterianos divididos de acordo com sua categoria.....	379
Figura 16: Projetos bacterianos organizados de acordo com sua filogenia.....	380
Figura 17: Porcentagem de projetos depositados no GOLD de acordo com seu gênero bacteriano.....	380

## LISTE DES FIGURES

Figure 1 : Aspect pléomorphe de <i>P. freudenreichii</i> .....	71
Figure 2 : La production du lait de vache sur le territoire français.....	76
Figure 3 : Processus de transformation des fromages PPC.....	79
Figure 4 : <i>P. freudenreichii</i> dans le fromage Emmental.....	80
Figure 5 : Schéma de réaction d'hydrolyse de triacylglycérides.....	80
Figure 6 : Représentation de la formule chimique de l'acide propionique.....	84
Figure 7 : Publications dans PubMed avec les mots clés « probiotics bacteria » au cours des dernières années.....	87
Figure 8 : Espèces microbiennes dominantes dans les parties du tractus gastro-intestinal.....	88
Figure 9 : Induction de l'apoptose à partir de SCFAs produites par les bactéries propioniques dans le côlon.....	90
Figure 10 : Combinaison des données génomiques, transcriptomiques et protéomiques.....	99
Figure 11 : Voie métabolique de Wood-Werkman chez <i>P. freudenreichii</i> liée au cycle de Krebs.....	100
Figure 12 : Schéma des processus métaboliques impliqués dans l'affinage des fromages PPC visant l'obtention des composés aromatiques.....	101
Figure 13 : Nombre total de projets par an déposés dans le GOLD.....	381
Figure 14 : Distribution des projets génomiques dans les bactéries selon leur classement.....	381
Figure 15 : La pertinence des projets bactériens présentés selon leur catégorie.....	381
Figure 16 : Projets bactériens organisé selon leur phylogénie.....	382
Figure 17 : Pourcentage des projets déposés dans GOLD selon leur genre bactérien...	382

## **LISTA DE TABELAS**

Tabela 1: Ação microbiana nos alimentos fermentados. ....	40
Tabela 2: Principais fermentadores para a indústria de laticínios.....	41
Tabela 3: Produção de queijos PPC em toneladas. ....	43
Tabela 4: Etapas para a fabricação de queijos. ....	44
Tabela 5: Genomas depositados de <i>P. freudenreichii</i> pela equipe do STLO. ....	64

## **LISTE DES TABLEAUX**

Tableau 1 : L'action microbienne dans les aliments laitiers fermentés.....	74
Tableau 2 : Principaux ferments pour l'industrie laitière et fromagère.....	75
Tableau 3 : La production des fromages PPC en tonnes.....	77
Tableau 4 : Les étapes de la fabrication des fromages.....	78
Tableau 5 : Génomes de <i>P. freudenreichii</i> déposés par le STLO.....	97

## LISTA DE ABREVIATURAS

	Inglês	Português
ATP	adenosine triphosphate	adenosina trifosfato
BLAST	basic local alignment search tool	ferramenta de busca de alinhamento local básico
bp	base pairs	pares de bases
cDNA	complementary DNA	DNA complementar
CFU	colony forming unit	unidade formadoras de colônias
CIRM-BIA	biological resource centre fully dedicated to bacteria of food interest	centro de recursos biológicos especializado em bactérias de interesse alimentar
DNA	deoxyribonucleic acid	ácido desoxirribonucleico
DHNA	1,4-dihydroxy-2-naphthoic acid	ácido 1,4-dihidróxi 2-naftóico
EMBL-EBI	European Molecular Biology Laboratory - European Bioinformatics Institute	laboratório europeu de biologia molecular - instituto europeu de bioinformática
FAO	food and agriculture organization	organização de alimentação e da agricultura
GOLD	genomes online database	banco de dados de genomas disponíveis ( <i>online</i> )
GRAS	generally recognised as safe	geralmente reconhecido como seguro
INRA	national institute of agronomic research	instituto nacional de pesquisa agrônômica
KEGG	kyoto encyclopedia of genes and genomes	enciclopédia kyoto de genes e genomas
LGCM	laboratory of cellular and molecular genetics	laboratório de genética celular e molecular
mRNA	messenger RNA	RNA mensageiro
NGS	next-generation sequencing	sequenciamento de próxima geração
OD	optical density	densidade óptica

ORFs	open reading frame	quadro aberto de leitura
PCR	polymerase chain reaction	reação em cadeia da polimerase
PPC	cooked pressed paste	massa prensada cozida
qPCR	real-time quantitative PCR	PCR quantitativa em tempo real
QPS	qualified presumption of safety	presunção qualificada de segurança
RNA-Seq	RNA sequencing	sequenciamento de cDNA
rRNA	ribosomal ribonucleic acid	ácido ribonucleico ribossômico
RT-PCR	reverse transcription - polymerase chain reaction	reverse transcriptase - réaction en chaîne par polymérase
SCFA	short-chain fatty acids	ácidos graxos de cadeia curta
STLO	science and technology of milk and eggs	ciência e tecnologia do leite e dos ovos
tRNA	transfer ribonucleic acid	ácido ribonucleico transportador
VBNC	viable but non-culturable	viáveis mas não cultiváveis
WHO	world health organization	organização mundial da saúde
YEL	yeast extract lactate	extrato de levedura-lactato

## LISTE DE ABREVIATIONS UTILISEES

	Anglais	Français
ATP	adenosine triphosphate	adénosine-5 -triphosphate
BLAST	basic local alignment search tool	outil de recherche de base d'alignement local
bp	base pairs	paires de bases
cDNA	complementary DNA	ADN complémentaire
CFU	colony forming unit	unité formant colonie (UFC)
CIRM-BIA	biological resource centre fully dedicated to bacteria of food interest	centre international de ressources microbiennes - bactéries d'intérêt alimentaire
DNA	deoxyribonucleic acid	acide désoxyribonucléique (ADN)
DHNA	1.4-dihydroxy-2-naphthoic acid	l'acide 1,4-dihydroxy -2-naphtoïque
EMBL-EBI	european molecular biology laboratory - european bioinformatics institute	laboratoire européen de biologie moléculaire - Institut européen de bio-informatique
FAO	food and agriculture organization	organisation des Nations Unies pour l'alimentation et l'agriculture
GOLD	genomes online database	base de données de génomes en ligne
GRAS	generally recognised as safe	présentant une innocuité sanitaire / généralement reconnu comme sûr
INRA	national institute for agricultural research	institut national de la recherche agronomique
KEGG	kyoto encyclopedia of genes and genomes	encyclopédie de Kyoto des gènes et des génomes
LGCM	laboratory of cellular and molecular genetics	laboratoire de génétique cellulaire et moléculaire
mRNA	messenger RNA	ARN messenger
NGS	next-generation sequencing	séquençage de prochaine génération
OD	optical density	densité optique

ORFs	open reading frame	cadre ouvert de lecture
PCR	polymerase chain reaction	réaction en chaîne de la polymérase
PPC	cooked pressed paste	pâte pressée cuite
qPCR	real-time quantitative PCR	PCR quantitative en temps réel
QPS	qualified presumption of safety	présomption d'innocuité reconnue
RNA-Seq	RNA sequencing	séquençage à ARN
rRNA	ribosomal ribonucleic acid	acide ribonucléique ribosomique
RT-PCR	reverse transcription - polymerase chain reaction	transcriptase reversa seguida da reação em cadeia da polimerase
SCFA	short-chain fatty acids	acides gras à chaîne courte
STLO	science and technology of milk and eggs	science & technologie du lait & de l'œuf
tRNA	transfer ribonucleic acid	acides ribonucléiques de transfert
VBNC	viable but non-culturable	viables mais non cultivables
WHO	world health organization	organisation mondiale de la santé
YEL	yeast extract lactate	bouillon lactate et extrait de levure



# **APRESENTAÇÃO / PRÉSENTATION**

## Colaborações

Este trabalho de tese foi realizado em Cotutela de acordo com as convenções propostas pelos estabelecimentos francês e brasileiro respectivamente, *Agrocampus Ouest*, representado pelo Sr. Grégoire Thomas e Universidade Federal de Minas Gerais – UFMG, representado pelo seu Reitor, na época, Professor Clélio Campolina Diniz. A tese foi realizada no laboratório de *Science & technologie du lait et de l'œuf* (STLO) localizado no *Agrocampus Ouest* (UMR 1253 - *Institut National de la Recherche Agronomique*) em Rennes (França) e no laboratório de Genética Celular e Molecular (LGCM) situado no Instituto de Ciências Biológicas na Universidade Federal de Minas Gerais (UFMG) em Belo Horizonte (Brasil). O desenvolvimento das atividades foi coordenado na França pelo Dr. Yves Le Loir e no Brasil pelo Prof. Dr. Vasco Azevedo. O convênio estabelecido em Cotutela permite a obtenção de um duplo diploma entre as instituições conveniadas, que são: Instituto Superior de Ciências Agrícolas, Agro-Alimentaires, Hortícolas e de Paisagem (*Agrocampus Ouest*), como representante francesa, e a UFMG, como representante brasileira.

Desde 1979, vários projetos foram realizados em colaboração franco-brasileira, envolvendo o Ministério da Educação (Brasil) através da CAPES, e o Ministério das Relações Exteriores e Européias e o Ministério do Ensino Superior e da Pesquisa (França), através do COFECUB. O financiamento da tese foi possível graças a este programa CAPES-COFECUB (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Comitê Francês de Avaliação da Cooperação Universitária com o Brasil) que permitiram o desenvolvimento deste doutorado durante 18 meses em Rennes (França). Este programa tem como objetivo, o intercâmbio científico de alunos da pós-graduação entre instituições de ensino superior do Brasil e da França, bem como visa a formação de recursos humanos nestes países.

## **Collaborations**

Cette thèse a été réalisée en co-tutelle selon les conventions proposées par les établissements français et brésilien, respectivement Agrocampus Ouest, représenté par M. Grégoire Thomas et Universidade Federal de Minas Gerais – UFMG, représenté par son recteur, à l'époque le Professeur Clélio Campolina Diniz. La thèse a été réalisée au laboratoire de Science & technologie du lait et de l'œuf (STLO) situé à Agrocampus Ouest (UMR 1253, à l'Institut National de la Recherche Agronomique) à Rennes (France) et au laboratoire de Génétique Cellulaire et Moléculaire situé dans l'Institut de Sciences Biologiques à l'UFMG à Belo Horizonte (Brésil). Le développement des activités a été coordonné en France par le Dr. Yves Le Loir et au Brésil par le Prof. Dr Vasco Azevedo. L'accord établi en co-tutelle permet d'obtenir un double diplôme entre les institutions partenaires que sont l'Institut Supérieur des Sciences Agronomiques, Agroalimentaires, Horticoles et du Paysage (Agrocampus Ouest) du côté français, et l'UFMG du côté brésilien.

Depuis 1979, plusieurs projets ont été réalisés dans le cadre de la coopération franco-brésilienne, avec le Ministère de l'Éducation (Brésil), par le biais de la CAPES, et du Ministère des Affaires Etrangères et Européennes et du Ministère de l'Enseignement Supérieur et de la Recherche (France), représenté par COFECUB. Le financement de la thèse a été possible grâce au programme CAPES-COFECUB (Coordination pour le Perfectionnement du Personnel de l'Enseignement Supérieur – Comité français d'évaluation de la coopération universitaire et scientifique avec le Brésil) qui a permis le développement de ce doctorat pendant 18 mois à Rennes (France). Ce programme vise l'échange scientifique des étudiants universitaires dans les établissements d'enseignement supérieur du Brésil et de France, ainsi que la formation des ressources humaines dans ces pays.

## Prefácio

No final do século XVI, bactérias foram descritas pela primeira vez por Antoine van Leeuwenhoek. Algumas bactérias foram exploradas à partir de experimentos realizados pelo Robert Koch, o qual as descreveu inicialmente como patógenos de animais e de homens, que viria a ser, mais tarde, denominado de *Mycobacterium pseudotuberculosis* (bacilo de Koch). Na mesma linha, trabalhos do Louis Pasteur, pesquisador francês, contribuíram para a compreensão das bactérias, seres procariotos e ubíquos, capazes de crescer em diferentes temperaturas, ambientes, intensidades de luz e presença e/ou ausência de oxigênio. Isto demonstra a capacidade de adaptação destes seres e sua aptidão de enfrentar condições favoráveis ou desfavoráveis. Um componente chave para esta aptidão é a proliferação e sobrevivência bacteriana por longos períodos (*Long-Term Survival*). Existem diferentes graus de sobrevivência desde a fase de latência de uma bactéria até o período denominado de “sobrevivência a longo prazo”. Apesar de muito utilizado na literatura, este conceito de “sobrevivência a longo prazo” está associado a persistência bacteriana em diferentes frações de tempo, como: horas, dias, meses e até anos. Esta sobrevivência a longo prazo é observada em bactérias isoladas de ambientes variados, tais como: solo, queijo, água, fósseis e de humanos (Greenblatt *et al.* 2004a; Lemunier *et al.* 2005; Finkel 2006a).

Dentre as bactérias que apresentam esta capacidade de sobreviver por longos períodos, estão aquelas pertencentes ao Filo Actinobacteria. Estudos de sobrevivência à longo prazo já foram realizados com diferentes espécies, como: *Micrococcus luteus* (Greenblatt *et al.* 2004a), *Mycobacterium bovis* (Fine *et al.* 2011), *M. tuberculosis* (Primm *et al.* 2000), *M. leprae* (Wheat *et al.* 2014), *Bifidobacterium lactis* (Olszewska, Staniewski and Łaniewska-Trokenheim 2012) e *Propionibacterium freudenreichii* (Dalmasso *et al.* 2012a). Dentro deste grupo bacteriano citado acima, temos bactérias que são classificadas como patogênicas e não-patogênicas. As bactérias patogênicas são mais estudadas, uma vez que são agentes de doenças e infecções de importância mundial. Uma pesquisa rápida no PubMed revela que termos como “pathogens bacteria survival” apresentam mais de 8400 resultados, enquanto que “probiotics bacteria survival” o retorno foi de 1074 resultados (acesso em <http://www.ncbi.nlm.nih.gov/pubmed> em 15-08-2015).

Para que bactérias patogênicas possam sobreviver é necessário que estas desenvolvam certas habilidades como aderir, invadir, colonizar o hospedeiro, evadir ao sistema imune e adquirir nutrientes a partir do hospedeiro, permitindo assim a adaptação

a diversos ambientes, muitas vezes hostis. Esse fenômeno de sobrevivência por longos períodos, já foi relatado em um modelo de infecção por *M. leprae*, o qual demonstrou que estas bactérias são capazes de reter a sua virulência *in vivo* (*nu/nu mouse model*) e estas permanecerem viáveis por até 8 meses dentro de cistos de ameba. Isto reforça a idéia de que estas bactérias podem persistir em vários hospedeiros ou ambientes (Wheat *et al.* 2014).

Em bactérias não-patogênicas, o estudo da sobrevivência apresenta um interesse biotecnológico, uma vez que as bactérias são utilizadas na agricultura, alimentação, medicina animal e humana. Na indústria alimentícia, sobretudo de laticínios, esta sobrevivência é essencial para manter uma atividade bacteriana nos produtos, por exemplo, probióticos que promovem funções benéficas ao hospedeiro (Lan *et al.* 2007).

Neste contexto, é interessante um bom entendimento sobre as estratégias de sobrevivência em bactérias não-patogênicas, principalmente de interesse biotecnológico na indústria alimentícia, como por exemplo, bactérias propiônicas. Estas últimas são alvo de pesquisa, desde o final dos anos 80, do laboratório de *Science & Technologie du Lait & de l'Oeuf (STLO)*, uma unidade mista de pesquisa situada no *Agrocampus Ouest* e INRA, e foi o instituto associado a este trabalho de Tese. Estudos recentemente realizados em *P. freudenreichii*, demonstraram que esta bactéria resiste a diferentes temperaturas, como nas etapas importantes de maturação dos queijos de massa prensada e cozida (PPC). Estas bactérias mantêm-se vivas em concentrações elevadas até o final da maturação e armazenamento destes produtos (Dalmasso *et al.* 2012a). Nos últimos anos, muitas pesquisas de biologia molecular têm sido desenvolvidas para a caracterização de *P. freudenreichii*, no intuito de melhorar a nossa compreensão a respeito da biologia desta bactéria, além de aprimorar a sua utilização em processos de fabricação de queijos ou em outras aplicações biotecnológicas (Parayre *et al.* 2007; Falentin *et al.* 2010c; Dalmasso *et al.* 2011). Além disso, *P. freudenreichii* é bom candidato probiótico, uma vez que ela sobrevive ao longo do trato gastrointestinal, exerce uma atividade imunomoduladora, produz substâncias antibacterianas e entra em competição com outras bactérias patogênicas por nutrientes e sítios de ligação (Jan, Rouault and Maubois 2000; Leverrier *et al.* 2003a; Lan *et al.* 2007).

Visando uma melhor compreensão da sobrevivência desta bactéria, o presente trabalho de tese aborda o estudo e a caracterização de estratégias de sobrevivência em *P. freudenreichii*.

O manuscrito é constituído em três capítulos: (I) revisão da literatura, (II) resultados e (III) discussão, conclusão e perspectivas. A primeira seção do trabalho é composta por uma revisão da literatura sobre o gênero *Propionibacterium* e sua espécie *P. freudenreichii*, auxiliando a nossa compreensão a respeito da biologia desta bactéria.

Ao longo dos últimos anos, o STLO já abordou o estudo da adaptação do *P. freudenreichii* a diferentes estresses ou ambientes, como em cólon de hospedeiro (modelo de porco) (Saraoui *et al.* 2013a), ou em processos de maturação de queijos PPC (Dalmasso *et al.* 2012a). Desse modo, os estudos da equipe STLO objetivaram a diferenciação das espécies de *P. freudenreichii* em subespécies, testando a capacidade para fermentação de lactose e redução de nitrato, além de analisarem a produção de compostos aromáticos. Isto permite selecionar uma linhagem de interesse na fabricação de queijos, além de ampliar conhecimentos gerais a cerca da fisiologia e metabolismo de *P. freudenreichii* (Thierry *et al.* 2011).

Assim, seguindo uma linha de pesquisa que visa melhor compreensão biológica deste microrganismo, o segundo capítulo deste manuscrito apresenta os resultados obtidos ao longo destes 4 anos de tese, sendo subdivido em duas seções: (I) caracterização fenotípica em diferentes fases do crescimento bacteriano durante um período de escassez de nutrientes de 11 dias; e (II) estudo da genômica funcional de *P. freudenreichii* para investigação ao nível molecular do fenômeno de adaptação à fase estacionária. Para isso, a metodologia utilizada foi o RNA-Seq, a qual permitiu analisar a expressão gênica global em fase exponencial e em fase estacionária de *P. freudenreichii*. Esta abordagem permitiu identificar os genes envolvidos na entrada da fase estacionária e previamente ao estado de dormência, em uma linhagem anteriormente selecionada por sua capacidade de sobrevivência à longo prazo. Como complemento a este estudo, os dados transcriptômicos foram associados aos dados bioquímicos (*in vitro*) gerados analisando a cinética do conteúdo das culturas em aminoácidos, ácidos orgânicos e açúcares, resultado do metabolismo desta espécie. Estas seções são destinadas a identificar as estratégias que permitem à *P. freudenreichii* de sobreviver à longo-prazo em condições de carência nutricional.

Como último capítulo, teremos a discussão geral deste trabalho, seguido de conclusões e perspectivas.

## Préambule

À la fin du XVII<sup>e</sup> siècle, les bactéries ont d'abord été décrites pour la première fois par Antoine van Leeuwenhoek. Certaines bactéries ont été explorées à partir d'expériences menées par Robert Koch, qui les décrit d'abord comme des pathogènes d'animaux et de l'homme, plus tard, ces bactéries recevront le nom de *Mycobacterium pseudotuberculosis* (bacille de Koch). Dans le même domaine, les travaux de Louis Pasteur contribuèrent à la compréhension des bactéries qui sont des êtres procaryotes et ubiquistes, capables de croître à différentes températures, dans différents environnements, à des intensités de lumière variables et en présence ou absence d'oxygène. Cela démontre la capacité d'adaptation de ces organismes et leur aptitude à affronter des conditions favorables ou défavorables. Une composante clé à cette aptitude est la prolifération et la survie bactérienne pendant de longues périodes (*Long-Term Survival*). Il existe différents degrés de survie dès la phase de latence d'une bactérie jusqu'à la période appelée « la survie à long terme ». Bien que cette expression soit très utilisée en littérature, ce concept de « survie à long terme » est associé à la persistance bactérienne à diverses échelles de temps, comme : l'heure, le jour, le mois et même l'année. Cette survie à long terme s'applique aux bactéries isolées d'environnements très variés, tels que: le sol, le fromage, l'eau, les fossiles et les êtres humains (Greenblatt *et al.* 2004a; Lemunier *et al.* 2005; Finkel 2006a).

Certaines des bactéries particulièrement reconnues pour leur capacité de survie à long terme appartiennent au *phylum* des *Actinobacteria*. Les études de survie à long terme ont été réalisées sur différentes espèces, telles que : *Micrococcus luteus* (Greenblatt *et al.* 2004a), *Mycobacterium bovis* (Fine *et al.* 2011), *M. tuberculosis* (Primm *et al.* 2000), *M. leprae* (Wheat *et al.* 2014), *Bifidobacterium lactis* (Olszewska, Staniewski and Łaniewska-Trokenheim 2012) et *Propionibacterium freudenreichii* (Dalmaso *et al.* 2012a). Dans le groupe bactérien mentionné ci-dessus, il y a aussi bien des bactéries pathogènes que non-pathogènes. Les bactéries pathogènes sont plus étudiées, car elles sont des agents de maladies et d'infections d'importance mondiale. Une recherche rapide sur PubMed révèle que des termes tels que « pathogens bacteria survival » présentent plus de 8400 entrées, tandis que « probiotics bacteria survival » en donnent 1074 (accès [www.ncbi.nlm.nih.gov/PubMed](http://www.ncbi.nlm.nih.gov/PubMed) en 15-08 -2015).

Pour que les bactéries pathogènes puissent survivre, il faut que celles-ci développent certaines capacités telles qu'adhérer, envahir, coloniser son hôte, échapper au système immunitaire et acquérir des nutriments à partir de l'hôte, leur permettant

ainsi de s'adapter à différents environnements, souvent hostiles. Ce phénomène de survie pendant de longues périodes a déjà été décrit dans un modèle d'infection pour *M. leprae*, qui démontre que ces bactéries sont capables de maintenir leur virulence, *in vivo*, sur modèle murin et qu'elles restent viables jusqu'à huit mois dans des kystes chez les amibes. Cela renforce l'idée que ces bactéries peuvent persister dans plusieurs hôtes voire plusieurs environnements (Wheat *et al.* 2014).

Dans les bactéries non-pathogènes, l'étude de la survie présente un intérêt biotechnologique, puisque que les bactéries sont utilisées dans l'agriculture, l'alimentation, la médecine vétérinaire et humaine. Dans l'industrie alimentaire, en particulier des produits laitiers, cette survie est essentielle pour maintenir une activité bactérienne dans les produits, par exemple, pour les probiotiques qui expriment des fonctions bénéfiques pour l'hôte (Lan *et al.* 2007).

Dans ce contexte, il est intéressant de bien comprendre les stratégies de survie des bactéries non-pathogènes, surtout d'intérêt biotechnologique dans l'industrie alimentaire, comme par exemple les bactéries propioniques. Depuis la fin des années 1980, celles-ci sont la cible de recherche du laboratoire de Science & Technologie du Lait et de l'Oeuf (STLO), une unité mixte de recherche (UMR) Agrocampus Ouest et INRA, où s'est déroulé ce travail de thèse. Des études récentes effectuées sur *P. freudenreichii* ont montré que cette bactérie est résistante à des températures variables, comme dans les étapes importantes d'affinage des fromages à pâte pressée cuite (PPC). Ces bactéries restent vivantes, à des concentrations élevées, jusqu'à la fin de l'affinage et pendant le stockage de ces produits (Dalmaso *et al.* 2012a). Au cours des dernières années, de nombreuses recherches de biologie moléculaire ont été mises au point pour la caractérisation de *P. freudenreichii*, afin d'améliorer notre compréhension biologique de cette bactérie, ainsi que pour améliorer son utilisation dans le processus de fabrication fromagère ou d'autres applications biotechnologiques (Parayre *et al.* 2007; Falentin *et al.* 2010c; Dalmaso *et al.* 2011). Par ailleurs, *P. freudenreichii* est une bonne candidate probiotique, car elle survit dans le tractus gastro-intestinal, exerce une activité immunomodulatrice, produit des substances antibactériennes et entre en compétition avec d'autres bactéries pathogènes pour les nutriments et les sites de liaison (Jan, Rouault and Maubois 2000; Leverrier *et al.* 2003a; Lan *et al.* 2007).

Pour améliorer notre compréhension de la survie de cette bactérie, ce travail de thèse aborde l'étude et la caractérisation des stratégies de survie chez *P. freudenreichii*.



Le manuscrit est constitué de 3 chapitres : (I) révision de la littérature, (II) les résultats et (III) la discussion, la conclusion et les perspectives. La première partie est composée d'une introduction bibliographique sur le genre *Propionibacterium* et l'espèce *P. freudenreichii* nous aider à comprendre le phénomène de survie à long terme chez *P. freudenreichii*.

Au cours de ces dernières années, le STLO a déjà abordé l'étude de l'adaptation de *P. freudenreichii* à différents stress ou dans différents environnements, tels que le côlon de l'hôte (modèle du porc) (Saraoui *et al.* 2013a), ou dans le processus d'affinage des fromages PPC (Dalmasso *et al.* 2012a). De même, les recherches de l'équipe STLO se sont focalisées sur la différenciation des espèces de *P. freudenreichii* en sous-espèces, en testant la capacité de fermentation de lactose et la réduction de nitrate, en plus d'analyser la production des composants aromatiques. Ceci permet de sélectionner une souche d'intérêt dans la fabrication des fromages, en plus d'élargir les connaissances générales liées à la physiologie et au métabolisme de *P. freudenreichii* (Thierry *et al.* 2011).

Ainsi, suivant une ligne de recherche qui vise une meilleure compréhension biologique de ce microorganisme, le deuxième chapitre de ce manuscrit présente les résultats obtenus au cours de ces quatre années de thèse, qui sont subdivisés en deux sections: (i) la caractérisation phénotypique à différentes phases de croissance bactérienne pendant une période de 11 jours de carence nutritionnelle; et (ii) l'étude de la génomique fonctionnelle de *P. freudenreichii* pour aborder au niveau moléculaire le phénomène d'adaptation à la phase stationnaire. Pour cela, la méthode utilisée a été le RNA-Seq, qui a permis l'analyse globale de l'expression des gènes en phase exponentielle et stationnaire de *P. freudenreichii*. Cela a permis d'identifier les gènes impliqués dans l'entrée en phase stationnaire et précédant l'état de dormance, dans une souche précédemment sélectionnée pour sa capacité de survie à long terme. En plus de cette étude, les données transcriptomiques ont été associées à des données biochimiques (*in vitro*) générées en analysant l'évolution du contenu des cultures en acides aminés, acides organiques et sucres résultant du métabolisme de cette espèce. Ces sections étaient destinées à identifier des stratégies qui permettent à *P. freudenreichii* de survivre à long terme en condition de carence nutritionnelle.

Le dernier chapitre, comprend une discussion générale de ce travail, suivie des conclusions et perspectives.

## **Objetivos / Objectifs**

## Objetivos

### Geral

Esta tese visa a identificar os mecanismos de sobrevivência a longo prazo em *Propionibacterium freudenreichii* que permitem sua adaptação à períodos prolongados durante a carência nutricional.

### Específicos

- ) Avaliar e caracterizar o fenótipo (morfologia, viabilidade e integridade de membrana) de 8 linhagens de *P. freudenreichii*, a fim de aferir a variabilidade entre as linhagens na capacidade de sobrevivência a longo prazo e de selecionar a linhagem que apresenta a melhor sobrevivência;
- ) Identificar a presença de células viáveis, mas não cultiváveis (VBNC) em *P. freudenreichii* através da associação de diferentes metodologias: UFC, microscopia de epifluorescência e qPCR;
- ) Identificar genes que são diferencialmente expressos em fases de crescimento exponencial e estacionária de *P. freudenreichii*;
- ) Analisar os genes possivelmente envolvidos no mecanismo de sobrevivência de *P. freudenreichii* de acordo com a técnica de RNA-Seq;
- ) Identificar, através de dados transcriptômicos e bioquímicos da linhagem selecionada (*P. freudenreichii* CIRM-BIA 138), quais são os genes presentes no organismo e quais são os fatores ambientais implicados para o desencadeamento de mecanismos moleculares que permitem a sobrevivência desta bactéria durante um período de carência nutricional;
- ) Correlacionar os resultados obtidos na análise do transcriptoma com os dados bioquímicos (dosagem de ácidos orgânicos, aminoácidos e açúcares) para validar a atividade metabólica de *P. freudenreichii*.

## Objectifs

### Généraux

Cette thèse vise à identifier les mécanismes de survie à long-terme chez *Propionibacterium freudenreichii* qui permettent l'adaptation à de longues périodes de carence nutritionnelle.

### Spécifiques

- Évaluer et caractériser le phénotype (la morphologie, la viabilité et l'intégrité membranaire) de 8 souches de *P. freudenreichii*, afin d'évaluer la variabilité inter-souches dans la capacité de survie à long terme et de sélectionner la souche présentant la meilleure survie du panel ;
- Identifier la présence de cellules viables mais non cultivables (VBNC) en *P. freudenreichii* par l'association de différentes méthodologies : UFC, microscopie épifluorescence et PCRq ;
- Identifier des gènes qui sont différentiellement exprimés dans les phases de croissance exponentielle et stationnaire de *P. freudenreichii* ;
- Analyser les gènes possiblement impliqués dans le mécanisme de survie à long-terme de *P. freudenreichii* selon la technique de RNA-Seq ;
- Identifier, à travers les données transcriptomiques et biochimiques de la souche sélectionnée (*P. freudenreichii* CIRM-BIA 138), quels sont les gènes présents et les facteurs environnementaux impliqués dans le déclenchement des mécanismes moléculaires permettant la survie de cette bactérie pendant une période de carence nutritionnelle ;
- Corréler les résultats obtenus dans l'analyse du transcriptome avec les données biochimiques (dosage d'acides organiques et aminés et des sucres) pour valider l'activité métabolique de *P. freudenreichii*.

**CAPÍTULO 1: REVISÃO DA  
LITERATURA**

**CHAPITRE 1 : RÉVISION DE LA  
LITTÉRATURE**

Esta parte do manuscrito é dedicada a apresentação de dados gerais sobre o gênero *Propionibacterium* e a apresentação, mais detalhada, da espécie *P. freudenreichii*, relativo a sua utilização na indústria, suas propriedades metabólicas e probióticas, e sua capacidade de resistir aos diferentes estresses.

## 1. Gênero *Propionibacterium*

### 1.1. Ecologia, fisiologia e classificação

Em 1906, Von Freudenreich e Orla-Jensen, relataram o primeiro isolamento de bactérias propiônicas a partir do queijo Emmental (von Freudenreich and Orla-Jensen 1906). Em 1909, o nome do gênero foi estabelecido por Orla-Jensen, sendo então denominado de *Propionibacterium* (Cummins and Johnson 1986), o qual foi proposto devido a capacidade desta bactéria de produzir ácido propiônico durante o processo de fermentação (Orla-Jensen 1909). Posteriormente, em 1946, Douglas e Gunter propuseram a inserção de bactérias pertencentes ao grupo de *Corynebacterium acnes* dentro do gênero de *Propionibacterium*. E, mais tarde em 1963, outros autores confirmaram esta reclassificação, sendo então *C. acnes* uma espécie incluída do gênero *Propionibacterium* (Moore and Cato 1963).

Bactérias do gênero *Propionibacterium* são bactérias Gram-positiva, pertencente à ordem Actinomycetales, dentro do filo Actinobacteria. Elas são imovéis e não formam esporos. Este gênero pode ser dividido em dois grandes grupos de acordo com o seu habitat: (i) lactícinos, os quais encontramos bactérias propiônicas lácticas; (ii) microbiota comensal da pele e/ou de membranas mucosas, os quais encontramos bactérias cutâneas ou comensais. No grupo de bactérias propiônicas lácticas, podemos encontrar diferentes espécies, as quais as mais conhecidas são: *P. freudenreichii*, *P. thoenii*, *P. jensenii* e *P. acidipropionici*. Estas espécies não são patogênicas para os seres humanos, e geralmente, são isoladas de produtos alimentares fermentados, como: silagem, fermentação de olivas, azeitonas e destilaria de rum. Em casos mais raros, podem ser isolados do meio-ambiente, por exemplo, em solos, e de animais, em especial no trato intestinal de ratos (Mantere-Alhonen 1982; Balows 1992). Outras espécies como: *P. cyclohexanicum* (Kusano *et al.* 1997) e *P. microaerophilum* (Koussémon *et al.* 2001) também foram recentemente descritas e classificadas pertencentes a este grupo de bactérias propiônicas lácticas não patogênicas, apesar de não serem isoladas de produtos lácteos, mas de alimentos variados ou de subprodutos. Exemplos destas ocorrências são:

*P. cyclohexanicum* encontrada em suco fermentado de laranja, e *P. microaerophilum* descrita em águas residuais de lagares de azeite. Entretanto, no grupo de *Propionibacterium* comensal ou cutânea, existem três espécies, tais como: *P. acnes*, *P. avidum* e *P. propionicum*, causando patologias em seres humanos. As espécies *P. lymphophilum* e *P. granulosum* pertencem também a classificação de comensais, porém representam um grupo distinto por estarem, filogeneticamente, mais distante de outros grupos (Meile *et al.* 1999).

## 1.2. Caracterização e importância do microrganismo em estudo: *Propionibacterium freudenreichii*

Em condições laboratoriais, o crescimento de *Propionibacterium freudenreichii* ocorre geralmente em condições ótimas no meio YEL (*Yeast Extract Lactate*) proposto por Malik e colaboradores (Malik, Reinbold and Vedamuthu 1968) e a uma temperatura de incubação de 30°C. Esta bactéria é anaeróbia facultativa que pode ser encontrada *in vitro* sob diversas formas (pleomórfica): cocóide, bífida ou ramificada (Figura 1).

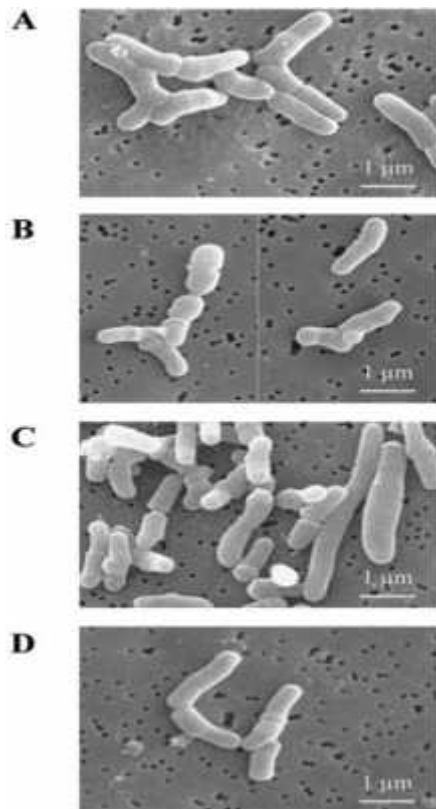


Figura 1: Aspecto pleomórfico de *P. freudenreichii* durante o estresse ácido: (A) células não tratadas; (B) células submetidas a pH= 2,0; (C) células submetidas a pH= 5,0; (D) células submetidas a pH= 5,0, e posteriormente, submetidas ao pH=2,0 visualizadas por microscopia eletrônica de varredura (Jan *et al.*, 2001).

Antigamente, a espécie *P. freudenreichii* foi dividida em duas sub-espécies: *P. freudenreichii* subsp. *freudenreichii*, que não é capaz de fermentar lactose, mas pode reduzir nitrato; e a *P. freudenreichii* subsp. *shermanii*, que fermenta lactose, porém não é capaz de reduzir o nitrato (Cummins and Johnson 1986). Porém, estudos recentes demonstram que existem linhagens de *P. freudenreichii* que podem conter os dois fenótipos: lactose e nitrato redutase positiva ou lactose e nitrato redutase negativa (de Carvalho, Gautier and Grimont 1994; Vorobjeva 1999; Dalmasso *et al.* 2011). Além disso, no genoma completo de *P. freudenreichii* CIRM-BIA 1, foi observado que o gene *lacZ*, responsável pela atividade beta-galactosidase, está localizado numa região de transposase e que existem pseudogenes na região do genoma que codifica a informação necessária à redução de nitrato (Falentin *et al.* 2010a). Atualmente, não há uma classificação de subespécies ao nível fenotípico e outras investigações para explorar a genômica desta bactéria devem ser efetuadas. Isto é grande importância na indústria alimentícia, uma vez que estas propriedades bioquímicas podem alterar a qualidade final do produto, por exemplo, queijos PPC.

*P. freudenreichii* é um microrganismo seguro (saúde-sanitária) e possui dois importantes status para sua utilização: o *Generally Recognized As Safe* (GRAS) outorgado pelo órgão administrativo de alimentos e drogas (*Food and Drug Administration – FDA, nos Estados Unidos*) (Mogensen *et al.* 2002) e o *Qualified Presumption of Safety* (QPS) concedido pela autoridade europeia de segurança dos alimentos (*European Food Safety Authority - EFSA*) (EFSA experts 2008). Estes status são aprovados por autoridades do setor que visam controlar a qualidade dos produtos, garantindo a segurança e a autorização destas bactérias na indústria de alimentos, como por exemplo, na fabricação de queijos Emmental. Estas bactérias são amplamente encontradas na natureza, incluindo o trato gastrointestinal, e estão presentes em muitos alimentos fermentados por possuírem características que as permitem de resistir ao pH ácido (Jan, Rouault and Maubois 2000; Huang and Adams 2004; Suomalainen *et al.* 2006) e a baixas temperaturas no processo de maturação de queijos (Dalmasso *et al.* 2012a). A concentração destas bactérias, geralmente, encontrada no Emmental são na ordem de  $10^9$  UFC/g. Estas bactérias possuem qualidades que são de grande importância para diversas aplicações na área industrial, as quais serão abordadas com mais detalhes na próxima seção.



## 2. Aplicações industriais de *Propionibacterium freudenreichii*

Os alimentos fermentados, em particular os laticínios, são consumidos pelo homem há milhões de anos. Desde o tempo da Idade Média, o processo de fabricação de queijos do tipo Emmental nos Alpes da região de Savoie é descrito em manuscritos. No século XIX, aparecem os primeiros relatos sobre “frutières”, que são locais de exploração e transformação do leite em queijo, onde os agricultores, em conjunto, agrupam o “fruto” de suas produções de verão. Em seguida, no século XX, uma forte industrialização foi observada, conduzindo a uma deslocalização da produção destes queijos.

Assim como o Emmental, todos os laticínios que foram anteriormente iniciados empiricamente, começaram a se beneficiar de avanços no campo da microbiologia e da tecnologia de alimentos, o que permitiu um melhor controle, diversificação e segurança de suas produções.

As principais espécies bacterianas utilizadas hoje para a produção de alimentos fermentados são pertencentes aos gêneros: *Bifidobacterium*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Streptococcus* e *Propionibacterium* (Tabela 1 e 2). Estes gêneros exercem papéis importantes na indústria de laticínios, na síntese de biomoléculas, bem como no desenvolvimento de produtos probióticos.

Neste contexto, *P. freudenreichii* desempenha um papel importante na fabricação e, particularmente, na maturação de queijos PPC. Esta espécie é também utilizada por sua capacidade de produzir bacteriocinas, compostos do ácido linoléico, do ácido propiônico, propileno, trealose e vitaminas do complexo B, como por exemplo, a vitamina B12 (Piao *et al.* 2004; Poonam *et al.* 2012). Algumas linhagens possuem também propriedades probióticas e podem modular a composição da microbiota intestinal produzindo fatores bifidogênicos, ou modular ao nível de inflamação no contexto de doenças inflamatórias intestinais crônicas. Estas bactérias ainda são interessantes por serem capazes de resistir ao estresse do trato gastrointestinal.

Como descrito anteriormente, esta bactéria apresenta capacidade de adaptação à estresses de natureza variável, como: pH, temperatura e sais biliares. Esta robusticidade confere uma vantagem a esta espécie que a permite de permanecer viva e de manter suas atividades metabólicas em diversos ambientes. Estas características são importantes para suas aplicações em indústrias alimentícias, particularmente, na indústria de queijos PPC; e em tecnologia, para a síntese de biomoléculas e/ou para

alimento/saúde (probióticos), todas estas aplicações serão apresentadas com detalhes nesta seção.

**Tabela 1: Ação microbiana nos alimentos fermentados (tabela adaptada de <http://mastervrv.free.fr/S3/BA/Cours/BA-FERMENTATION.pdf>).**

Efeitos sob o produto	Exemplos	Microrganismos implicados
Modificação da textura	Coagulação do leite	Diversas bactérias lácticas
	Textura aveludada em laticínios fermentados	Bactérias lácticas (dextranes)
	Formação de olhaduras em queijos	Bactérias propiônicas
Aromatização	Gosto amanteigado (diacetil)	<i>Lactococcus lactis diacetylactis</i>
	Gosto típico de iogurte (acetaldeído e ácido láctico)	<i>Lactobacillus delbrueckii bulgaricus</i>
	Desenvolvimento do gosto de cacau	Diversas bactérias
Coloração	Escurecimento da massa de queijos PPC* por propionato de cálcio	Bactérias propiônicas
Alteração do gosto	Acidificação do repolho	<i>Lactobacillus</i>
	Desacidificação do vinho	Bactérias malolácticas
Estabilização microbiana	Acidificação	Bactérias lácticas e acéticas

\*PPC: queijos de massa cozida e prensada.

**Tabela 2: Principais fermentadores para a indústria de laticínios (tabela adaptada de <http://mastervrv.free.fr/S3/BA/Cours/BA-FERMENTATION.pdf>).**

Fermentadores	Aplicações
<i>Lactobacillus bulgaricus</i> <i>Streptococcus thermophilus</i> <i>Bifidobacterium sp.</i>	Iogurtes
<i>Lactobacillus acidophilus</i>	Leite fermentado
<i>Lactobacillus caucasicus</i>	Kéfir (bebida fermentada)
<i>Streptococcus cremosnis</i> <i>Streptococcus lactis</i> <i>Streptococcus diacetylactis</i> <i>Leuconostoc cremonis</i> <i>Lactobacillus cremonis</i> <i>Lactobacillus lactis</i>	Etapas iniciais (coalhamento) da maioria dos queijos Queijos não afinados (massa fresca) Manteiga
<i>Propionibacterium freudenreichii</i> ssp. <i>shermanii</i> <i>Propionibacterium freudenreichii</i> ssp. <i>freudenreichii</i>	Maturação de queijos PPC*

\*PPC: queijos de massa cozida e prensada.

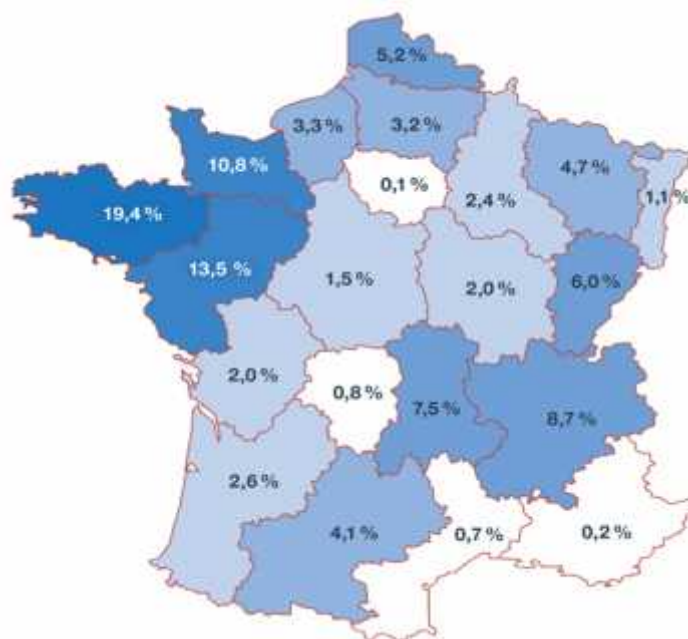
Nos parágrafos seguintes, iremos apresentar algumas informações sobre a importância de *P. freudenreichii* em aplicações industriais:

- ) **Queijos PPC** → as bactérias propiônicas são necessárias para a maturação dos queijos PPC. Sua sobrevivência é necessária para garantir o sabor e a tipicidade destes queijos.
- ) **Biomoléculas** → bactérias propiônicas permitem a produção de diversas biomoléculas de interesse industrial, como por exemplo, o ácido propiônico e a vitamina B12, produtos obtidos durante o metabolismo destas bactérias.
- ) **Probióticos** → a manutenção da atividade metabólica de *P. freudenreichii* nos produtos probióticos é de grande importância, uma vez que esta bactéria pode atuar no trato gastrointestinal, como por exemplo, através da imunomodulação. Para isto, é necessário que as bactérias sobrevivam por longos períodos passando por etapas de transformação e armazenamento, sem grandes alterações na viabilidade e atividade esperada para o produto final.

## 2.1. Indústria de queijos PPC (“pasta cozida e prensada”)

Os queijos são mundialmente produzidos e são classificados de acordo com suas características, tais como: consistência da pasta, teor de gordura, tipo de leite utilizado, tipo de coagulação, tipo de casca e tempo de cura. *P. freudenreichii* é utilizada em conjunto com outras culturas bacterianas, principalmente, na fabricação de queijos Emmental. Este grupo é o terceiro segmento mais importante na indústria de queijos, sendo chamado de “pasta cozida e prensada” (PPC, do francês: pâte pressée cuite), e composto por outros queijos, como o queijo Comté. A denominação PPC corresponde aos queijos, os quais a coagulação e desidratação são extensas e que, em seguida, são prensados no momento da moldagem para remover a quantidade máxima de soro do leite, nomeados também de “queijos duro”. Após a etapa de coagulação e moldagem, a coalhada é “cozida” em uma etapa de alto aquecimento, acima de 50°C durante 30 à 60 minutos no soro, o que permite de obter, após a desidratação e prensagem, um elevado extrato seco e uma conservação destes queijos por mais tempo.

De acordo com o *Centre National Interprofessionnel de l'Economie Laitière*, em 2014, a França produziu mais de 260.000 toneladas de Emmental. A região do *Grand Ouest* da França, contemplando a Bretanha, País de Loire e Baixa Normandia, representa dois terços desta produção (CNIEL, 2014). Destaque para a região da Bretanha que é a primeira em produção de leite de vaca et de Emmental na França (Figura 2).



**Figura 2: Produção de leite de vaca no território francês** (*La géographie du lait em 2013 – Enquête Annuelle Laitière*).

Estes produtos são muito importantes para a economia de lactícínios da França, particularmente o Emmental, que é um grande representante deste setor (Tabela 3).

**Tabela 3: Produção de queijos PPC em toneladas (L'Enquête Annuelle Latière).**

Queijos PPC	2011	2012	2013
Emmental	256.980	253.225	252.892
Comté AOP	58.114	57.561	59.459
Beaufort AOP	5.044	4.919	4.787
Gruyère IGP	-	2.679	2.929
Autres PPC	-	8.913	7.938
Total PPC	333.258	327.298	328.006

Para compreender o processo de fabricação destes queijos PPC, a tabela 4 resume algumas etapas importantes e a figura 3, ilustra todo o processo.

**Tabela 4: Etapas para a fabricação de queijos.**

<b>Etapas</b>	<b>Descrição</b>
<b>Pasteurização</b>	Aquecimento do leite a temperaturas elevadas, esta etapa visa a eliminação de microorganismos indesejáveis. Porém, esta etapa não ocorre para queijos feito com leite cru ou pasteurizado.
<b>Coagulação</b>	Esta etapa visa a concentração da proteína do leite, podendo ser dividida em duas etapas: (I) ácida e (II) enzimática. A primeira ocorre pela atividade da flora microbiana do leite, ou pela adição do extrato de renina (coalho) ou fermento. Já a coagulação ácida ocorre a partir da produção de ácido láctico por bactérias, ou pela adição de ácidos orgânicos ao leite.
<b>Moldagem</b>	A coalhada obtida no processo anterior é colocada em diferentes moldes, conforme o queijo desejado.
<b>Desidratação</b>	O objetivo desta etapa é separar o soro de leite da coalhada, o que prolonga sua conservação. Alguns queijos podem ser consumidos após este processo.
<b>Salga</b>	Os queijos frescos resultantes são desenformados e salgados, com um sal fino ou no banho de salmoura saturada. Isto garante uma ação anti-séptica, melhoria na conservação e também adiciona o sabor aos queijos.
<b>Maturação</b>	Esta final pode demorar dias ou até meses de acordo com a variedade do queijo. Estes serão armazenados em “caves”, local o qual terão a temperatura e umidade controlada.

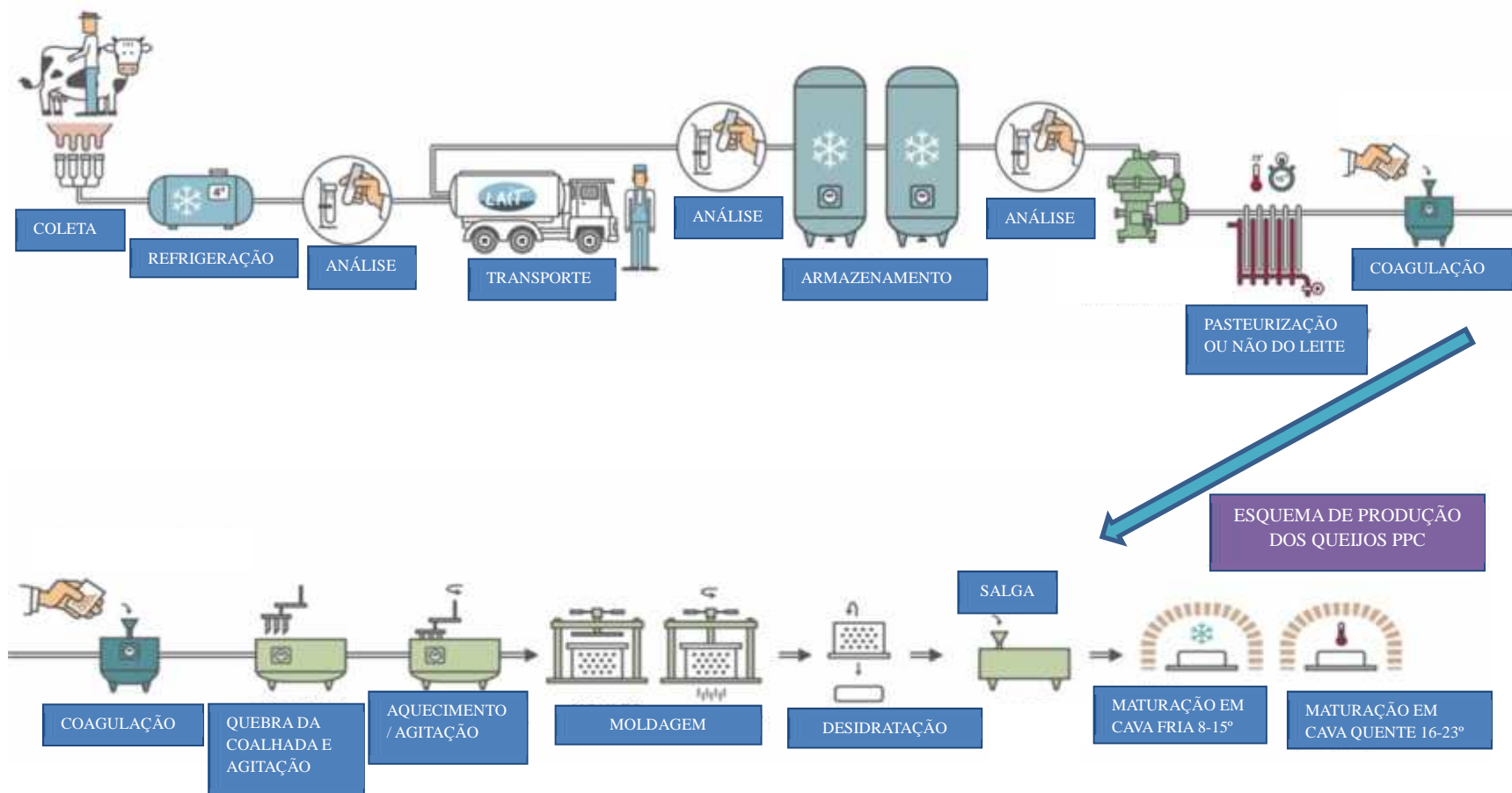


Figura 1: Processos de transformação de queijos PPC (<http://www.produits-laitiers.com/>). O processo de fabricação dos queijos PPC, inicia-se pela coleta do leite por meio de máquinas automatizadas em fazendas de gado leiteiro. Em seguida, este leite é refrigerado e analisado, de acordo com as normas de segurança do consumidor, visando garantir um produto final de boa qualidade. Este leite é transportado em caminhões isotérmicos até a indústria de laticínios, onde começam todas as etapas de tratamento e fabricação de queijos.

O Emmental é um queijo de peso, cujo as unidades pesam entre 60 e 120 kg (grande em volume de peso). É necessário 12 litros de leite para obter 1kg deste queijo (CNIEL,2015). Durante o processo de maturação, muitos parâmetros variam, tais como: temperatura, pH e atividade de água. *P. freudenreichii* exerce um papel importante nesta maturação (Figura 4), produzindo acetato e propionato através da fermentação, que estão envolvidos no sabor do Emmental, e CO<sub>2</sub> responsável pela “abertura do queijo” por meio da formação de olhaduras na massa (Langsrud and Reinbold 1973; Meurice *et al.* 2004).

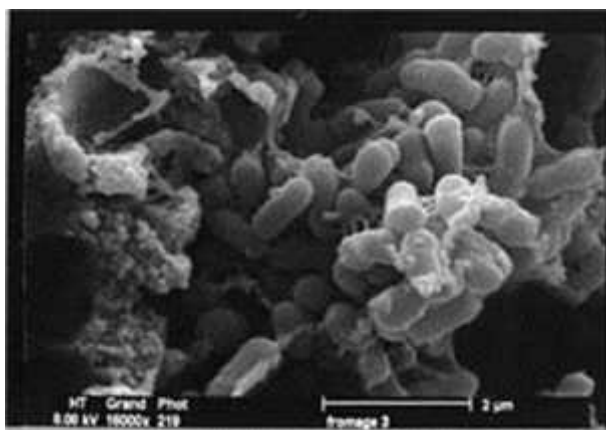


Figura 4: *P. freudenreichii* em um queijo Emmental (©INRA).

*P. freudenreichii* desempenha também um papel importante na composição do aroma e sabor em alguns laticínios, dentre eles o Emmental. Isto porque durante o processo de maturação prolongada deste queijo, *P. freudenreichii* garante a ocorrência das reações de lipólise. A lipólise contribui para a formação de sabores e é realizada através de esterases, as quais são capazes de efetuar a hidrólise dos triacilglicerídeos do leite liberando ácidos graxos de cadeia curta, de cadeia média e de cadeia longa (Figura 5). Estes últimos é que conferem aos queijos suas diferentes sensações gustatórias, como por exemplo, sabor forte, picante e/ou salgado (Berdagué *et al.* 1990).

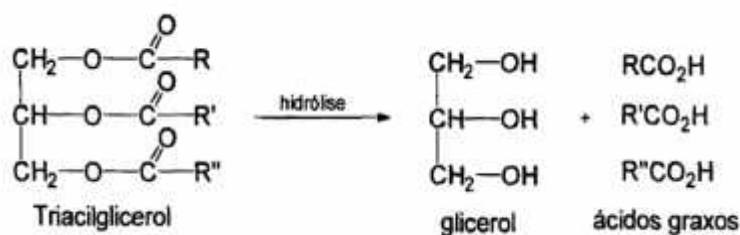


Figura 5: Esquema da reação de hidrólise de triacilglicerídeos.



Estudos como o de Dherbécourt e colaboradores, identificaram 23 esterases (*putative esterases*) em *P. freudenreichii*, potencialmente envolvidas neste processo de maturação. Esta identificação foi realizada a partir da combinação de métodos clássicos e do estudo da genômica desta bactéria (Dherbécourt *et al.* 2008). Através da abordagem genômica, as ferramentas bioinformáticas e os métodos utilizados foram: (i) busca automática de esterases por meio de homologia, (ii) busca por motivos protéicos no InterProScan, (iii) busca manual por homologia e (iv) busca manual por motivos específicos através do Prosite. Estes métodos foram essenciais para uma primeira triagem de enzimas de interesse. Assim, 12 genes candidatos que codificam as potenciais esterases apresentando os motivos GX SXG característicos de / hidrolase foram selecionados e clonados em *E. coli*. Testes zimográficos também foram realizados no intuito de caracterizar a atividade destas esterases. Assim, foi constatado que as 12 enzimas, em *P. freudenreichii*, desempenham papéis importantes na lipólise no queijo Emmental.

Sabemos que *P. freudenreichii* sobrevive a diversos ambientes e estresses de maneira linhagem-dependente. A importância deste fenômeno atrai estudos sobre as estratégias que permitem a sobrevivência e manutenção da atividade metabólica por longos períodos. Assim, no intuito de explorar estes mecanismos de sobrevivência à longo prazo, estudos que associam experimentos microbiológicos (*in vitro*) relacionado com aplicações industriais e biotecnológicas são de uma grande importância. Estes estudos serão descritos nas próximas seções.

## 2.2. Otimização da produção de vitamina B12

Em 1911, o termo “vitamina” foi proposto por Casimur Funk para designar as aminas que são essências para a vida (vita-amina). Posteriormente, esta nomenclatura começou a ser utilizada também para substâncias orgânicas essenciais, sem valor energético mas ativas em pequenas quantidades, contribuindo para a manutenção do metabolismo de um organismo (Florent 1986; Vandamme 1992).

A vitamina B12, também denominada de cobalamina ou cianocobalamina, foi isolada com primazia em 1948, tendo sido sintetizada somente em 1973. Isto ocorreu devida a dificuldade de obtenção do produto, e sua síntese necessita de diversas etapas de processamento até o rendimento final da molécula (Vandamme 1989). Como esta vitamina está implicada em processos essenciais para a regulação do metabolismo de animais e seres humanos, como a formação e maturação de hemácias e a manutenção do sistema nervoso, estratégias alternativas que visam a otimização de sua produção são de grande interesse.

Neste contexto, a obtenção de vitamina B12 utilizando microorganismos tem se tornado uma solução promissora para o mercado industrial. Dentre as bactérias produtoras, destaca-se a *P. freudenreichii*, devido ao seu bom rendimento de vitamina B12. Assim, uma concentração de 5,12 mg/L pode ser obtida no queijo após 168h de fabricação (Marwaha, Sethi and Kennedy 1983). Estudos visando aprimorar a produção de vitamina B12 em *P. freudenreichii* mostraram que é possível alcançar um rendimento de até 6,35 mg/L após 192h de fermentação (García Hernández and Pérez Mendoza 1983).

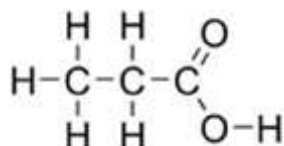
Novas abordagens tem sido propostas para a biossíntese de vitamina B12 ao decorrer das últimas décadas. A obtenção de vitamina B12 por meio da fermentação de *P. freudenreichii* foi analisada, no trabalho de Paio e colaboradores, de acordo com o seu desempenho sob a expressão de genes, como: *cobA*, *cbiLF* e *cbiEGH*, envolvidos na síntese de vitamina B12, poderia aumentar a produção de vitamina B12 (Piao *et al.* 2004). Outros estudos com o mesmo objetivo, como o trabalho de Murooka e colaboradores, expressaram e amplificaram os genes do operon *cob* e *cbi*, para o aumento da biossíntese de cobalaminas durante o processo de fermentação (Murooka *et al.* 2005).

Para otimizar a produção de vitamina B12, um outro estudo por Wang e colaboradores com enfoque no efeito do ácido propiónico e do 5,6-dimetilbenzimidazole durante a biossíntese desta vitamina no processo de cultivo de *P.*

*freudenreichii* (Wang *et al.* 2015). Foi visto que a concentração de ácido propiônico varia da fase inicial (10-20 g/L) à fase tardia (20-30 g/L). Isto conduz a um feed-back negativo para a biossíntese de vitamina B12, mas que pode ser alterado com a inclusão de 5,6-dimetilbenzimidazole (Wang *et al.* 2015). Assim, um controle e uma adaptação detalhada das condições de cultivo bacteriano são essenciais para a obtenção de grandes quantidades de vitamina B12.

### 2.3. Obtenção do ácido propiônico

O ácido propiônico é um ácido graxo de cadeia curta, saturado, de massa molecular 74,08 g/mol, também conhecido como ácido propílico, ácido propanóico ou  $C_3H_6O_2$  (GESTIS Substance Database – disponível em [http://gestis-en.itrust.de/nxt/gateway.dll/gestis\\_en/012590.xml?f=templates\\$fn=default.htm\\$3.0](http://gestis-en.itrust.de/nxt/gateway.dll/gestis_en/012590.xml?f=templates$fn=default.htm$3.0)) (Figura 6).



**Figura 6: Representação da fórmula química do ácido propiônico.**

Este ácido é amplamente utilizado em setores industriais, como: petroquímicos, cosméticos, plásticos, herbicidas e conservantes. Neste contexto, a biologia sintética é um setor inovador que envolve a biologia e a engenharia, com o intuito de desenvolver novas funções e sistemas, por exemplo, a utilização de uma determinada via bioquímica para a obtenção de um dado material por meio de uma fonte renovável (Parizzi *et al.* 2012).

A fermentação é um dos processos mais utilizados na indústria para a obtenção de diversos produtos e subprodutos, entre eles, o ácido propiônico. Atualmente, estratégias alternativas visando a obtenção deste produto foram desenvolvidas utilizando diferentes fontes de carbono, tais como a fermentação da cana-de-açúcar (sacarose), estas estratégias buscam uma redução no custo e uma melhoria na produção deste ácido. As bactérias propiônicas são capazes de sintetizar o ácido propiônico e elas representam uma fonte interessante para tal utilidade. Assim, estudos que visam a fisiologia e metabolismo de Propionibacterias foram realizados, dentre outros, no intuito de identificar as melhores espécies e as melhores linhagens para uma utilização industrial (Thierry *et al.* 2011).

A fermentação de lactato, através do metabolismo inicial de *P. freudenreichii*, gera produtos como acetato, propionato e  $CO_2$ , que desempenham papéis variados e importantes para as indústrias alimentares e biotecnológicas (Meurice *et al.* 2004). Por exemplo, o propionato de cálcio é um produto oriundo do ácido propiônico, tendo a capacidade de atuar como conservante, sendo capaz de combater bactérias desfavoráveis e fungos em alimentos.

A partir destas observações sobre o ácido propiônico que Stowers e colaboradores avaliaram 17 linhagens do gênero *Propionibacterium* e definiram parâmetros otimais de produção e de rendimento de ácido propiônico em fermentador. Com base nestes parâmetros, linhagens das espécies de *P. acidipropionici* e *P. freudenreichii* produziram altas quantidades de ácido propiônico (Stowers, Cox and Rodriguez 2014).

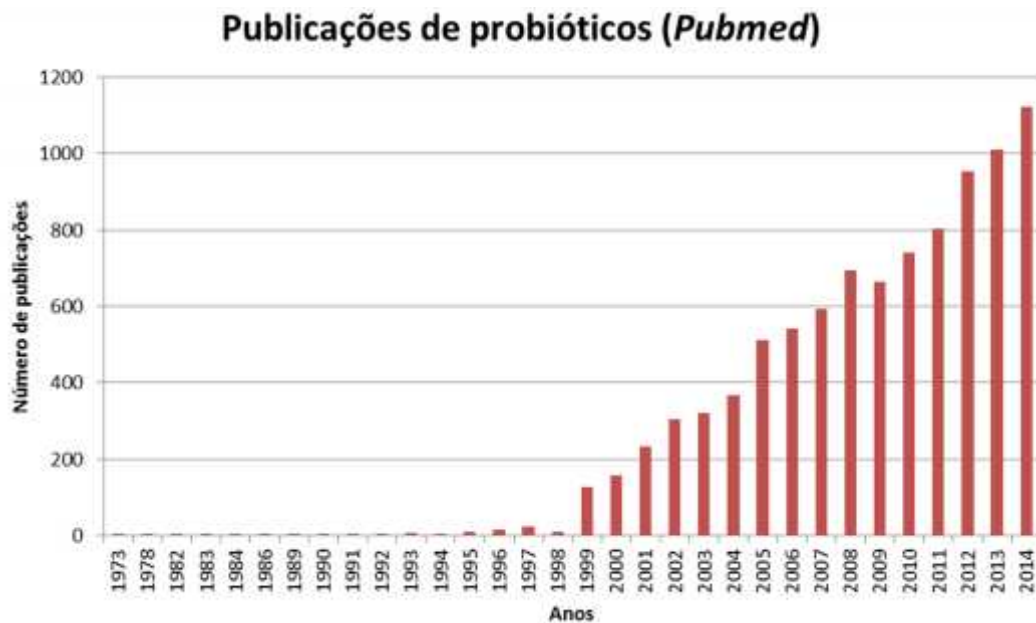
Porém, ainda é difícil a obtenção de um bom rendimento de ácido propiônico a partir de *Propionibacteria*, uma vez que a geração de subprodutos do seu metabolismo, como ácido succínico, ácido láctico, e ácido acético diminui o rendimento do ácido propiônico (Stowers, Cox and Rodriguez 2014). Neste intuito, é importante a realização de estudos para aumentar a eficácia da fabricação deste produto. Esta produção lança oportunidades em setores como o de química ambiental, através da produção de propileno ( $C_3H_6$ ), precursor de ácido propiônico, um dos hidrocarbonetos utilizados na indústria petroquímica. A polimerização do propileno forma o polipropileno, um composto muito utilizado em diversas indústrias para gerar combustíveis, embalagens descartáveis, produtos têxteis, eletrodomésticos e mercado automobilístico (Ammar *et al.* 2014). Assim, bactérias que sintetizam o ácido propiônico a um alto rendimento são objetos de interesse crescente, devido ao desenvolvimento da química ambiental e de um mercado econômico favorável para produtos orgânicos.

## **2.4. Probiótico**

Como falado anteriormente, diversas bactérias, em especial bactérias lácticas e propiônicas, tem sido alvo de interesse devido as suas propriedades probióticas. Um exemplo disso é *Propionibacterium freudenreichii* que é amplamente estudada por sua capacidade de inibir microrganismos indesejáveis, modular a microbiota do cólon através da produção de 1,4-di-hidroxi-2-naftóico (DHNA), um fator bifidogênico (Pérez Chaia, Zarate and Oliver 1999), e à sobreviver a diferentes estresses (Isawa *et al.* 2002; Okada *et al.* 2013), em particular aos estresses encontrados durante a passagem pelo trato gastrointestinal (Jan *et al.* 2001; Leverrier *et al.* 2003a; Saraoui *et al.* 2013a). *P. freudenreichii* é também investigada por desencadear a apoptose em células cancerígenas do cólon, aderir à células epiteliais do intestino e mucosa, estimulando assim atividades imunomodulatórias e a proliferação celular no hospedeiro (Ouwehand *et al.* 2002; Zarate, Chaia and Oliver 2002). Nesta seção, iremos apresentar uma pequena síntese de estudos que abordam as diversas propriedades probióticas atribuídas à *P. freudenreichii*.

### **2.4.1. Histórico e definição do termo probiótico**

Embora os benefícios de alguns alimentos fermentados para a saúde sejam estudados há muitos anos (trabalhos de Metchnikoff no início do século XX), o termo probiótico teve origem em 1965, quando Lilly e Stillwell, relacionaram a sobrevivência de determinados microrganismos com o efeito benéfico de seus produtos (Lilly and Stillwell 1965). Atualmente, os probióticos são definidos como microrganismos vivos que se administrados de maneira adequada, promovem benefícios à saúde do hospedeiro (Bron, van Baarlen and Kleerebezem 2012). Devida a sua importância, publicações no PubMed vem ganhando destaque ao longo dos últimos anos (Figura 7).



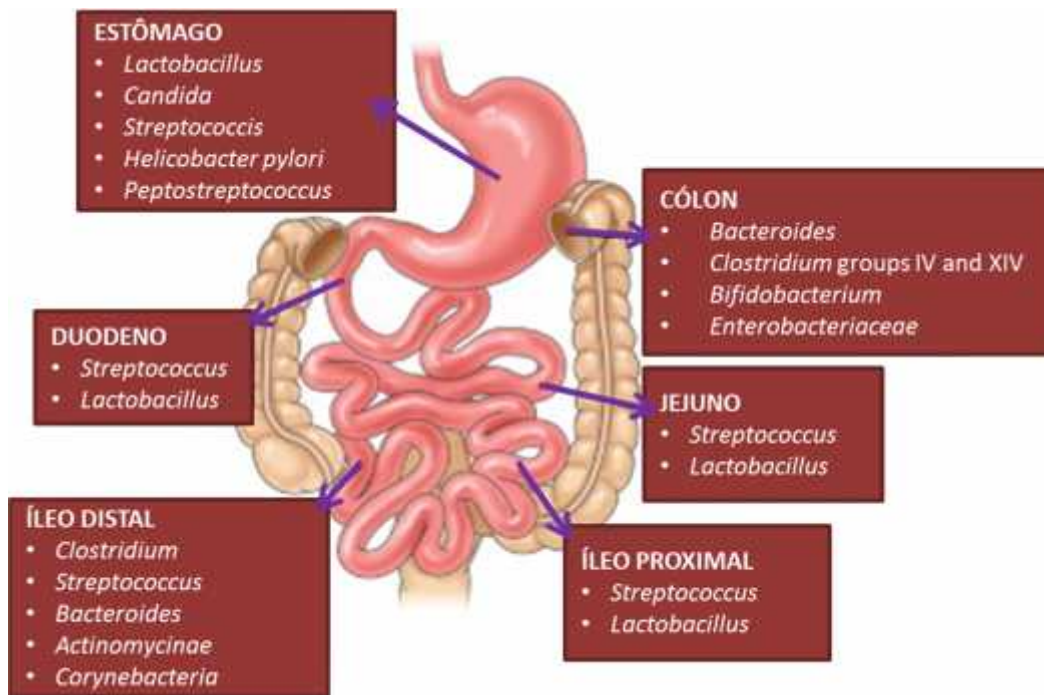
**Figura 7:** Publicações no PubMed com as palavras-chave “probiotics-bacteria” ao decorrer dos últimos anos.

Além disso, para a utilização de probióticos em humanos e para que as reivindicações de saúde sejam aceitas, são necessários rigorosos testes *in vitro* e *in vivo*, os quais visam garantir a qualidade e segurança destes microrganismos antes de suas possíveis aplicações tanto na prevenção ou no tratamento de enfermidades quanto em suplementos alimentares. Algumas espécies bacterianas como *Lactobacillus paracasei*, *L. johnsonii*, *L. fermentum*, *L. plantarum*, *L. rhamnosus*, *Bifidobacterium longum* e *P. freudenreichii* são objetos de estudo por apresentarem capacidades que promovem a melhoria da saúde humana, incluindo o controle e a restauração do equilíbrio da microbiota intestinal (Parvez *et al.* 2006; Cousin *et al.* 2012; Santos Rocha *et al.* 2012; Luerce *et al.* 2014).

#### 2.4.2. Controle da microbiota

Bactérias propiônicas podem exercer efeitos na microbiota dos animais e/ou de seres humanos, seja inibindo microrganismos não-desejáveis e/ou aumento a proliferação de populações de *Bifidobacterium* sp., bactérias que exercem funções importantes no trato digestivo (Picard *et al.* 2005). O seu aumento apresenta uma série de vantagens fisiológicas e clínicas para a saúde humana e/ou animal, tais como: (i) acelera o tempo de passagem pelo trato gastrointestinal; (ii) exerce o efeito de barreira, já que produz substâncias que bloqueiam os sítios de ligação de patógenos; (iii) exerce efeito sobre o sistema imunológico e em doenças do trato-gastrointestinal; (iv) reduz o

risco de diarreia por infecção por rotavírus; (v) produz nitrosaminas, nitritos e aminas heterocíclicas; (vi) exerce efeitos sobre a proliferação de células intestinais e (vii) atua na carcinogênese do cólon (Picard *et al.* 2005). Além de Bifidobactérias, existem outras espécies bacterianas dominantes que colonizam diversas porções do trato gastrointestinal e que são de grande importância para o controle da microbiota intestinal. Estas espécies são mostradas na Figura 8.



**Figura 8: Espécies microbianas dominantes em porções do trato gastrointestinal. Figura adaptada do artigo “Microbial Influences in Inflammatory Bowel Diseases” (Sartor 2008).**

A microbiota intestinal exerce um papel importante no sistema imunológico dos seres humanos e/ou animais e esta é, geralmente, adquirida durante o primeiro ano de vida logo após ao nascimento. A microbiota pode ser influenciada por fatores ambientais externos, como por exemplo, a ingestão de produtos lácteos na dieta (Bernstein 2014; Geuking *et al.* 2014).

A composição da microbiota intestinal apresenta um interesse fundamental quando associada a patologias do trato gastrointestinal. Diversos estudos buscam também um melhor entendimento destas comunidades microbianas e seus efeitos de modulação do sistema imune e, em geral, da saúde do hospedeiro (Sartor 2008; Hooper and Macpherson 2010; de Vos and de Vos 2012; Vangay *et al.* 2015; Velasquez-Manoff 2015).



### 2.4.3. Imunomodulação

A imunomodulação é uma propriedade importante de bactérias probióticas, permitindo a indução e regulação de respostas imunes, sendo útil na prevenção e no tratamento de câncer, de doenças auto-imunes e no controle de microrganismos indesejáveis (Ouwehand *et al.* 2009; Sathish *et al.* 2013).

A associação de bactérias propiônicas à bactérias lácticas e bifidobactérias podem exercer efeitos benéficos em patologias do trato gastrointestinal em crianças e idosos. Neste contexto, bactérias propiônicas agem indiretamente favorecendo o crescimento e a manutenção de concentrações elevadas de bifidobactérias durante um período de tratamento (Mantere-Alhonen 1995). As bactérias propiônicas parecem somente promover um efeito probiótico através da imunomodulação (Cousin *et al.* 2010; Le Marechal *et al.* 2015). O efeito imunomodulador de linhagens probióticas é comumente associado a sua capacidade de adesão ao epitélio intestinal. Esta adesão permite a permanência destas bactérias por mais tempo no trato digestivo, durante o trânsito intestinal. Esta capacidade de adesão pode ser um critério de seleção de probióticos e algumas linhagens de *P. freudenreichii* desencadearam uma série de estudos neste campo (Ouwehand *et al.* 2002). A implicação de proteínas de superfície no papel imunomodulatório destas bactérias foi testado através da caracterização proteômica “da superfície” de *P. freudenreichii*. Foi demonstrado também que algumas proteínas de superfície, extraídas de linhagens de *P. freudenreichii* e purificadas, foram capazes de reproduzir, *in vitro*, uma resposta anti-inflamatória observada com uma linhagem completa (Le Maréchal *et al.* 2015).

#### 2.4.4. Apoptose induzida por ácidos graxos de cadeia curta

As bactérias propiônicas são conhecidas por sua capacidade de produzir ácidos graxos de cadeia curta (SCFA – short chain fatty acids), entre eles, o propionato e o acetato que são os principais produtos da fermentação de *P. freudenreichii*. Foi observado que estes SCFA apresentam efeitos interessantes sob a fisiologia das células epiteliais intestinais (Jan *et al.* 2002; Pérez Chaia and Zarate 2005).

Os SCFA, e principalmente o butirato, constituem uma fonte de energia privilegiada, um verdadeiro “combustível” para as células epiteliais intestinais. Em 1995, Hague e colaboradores demonstraram que estes mesmos SCFA, por exemplo, o butirato de sódio, são também capazes de induzir a apoptose em cancerígenas do cólon (Hague *et al.* 1995). Posteriormente, o trabalho de Jan e colaboradores em 2002, demonstra que *in vitro* as linhagens de *P. acidipropionici* e *P. freudenreichii* são também capazes de induzir a apoptose em linhagens de células humanas de câncer colorretal, como: HT29 e Caco2. Este efeito também foi observado em linhagens celulares humanas de câncer de cólo de útero (HeLa). Os SCFA induzem sinais característicos de apoptose, tais que: (i) perda do potencial transmembranar das mitocôndrias, (ii) geração de espécies reativas de oxigênio, (iii) indução da cascata caspase-3 e (iv) condensação da cromatina ao nível nuclear. O mecanismo geral de apoptose através de SCFA é esquematizada na Figura 9. Além disso, esta atividade pode ser exercida *in vivo*, já que as linhagens se mantêm vivas e metabolicamente ativas dentro do intestino humano (Jan *et al.* 2002).

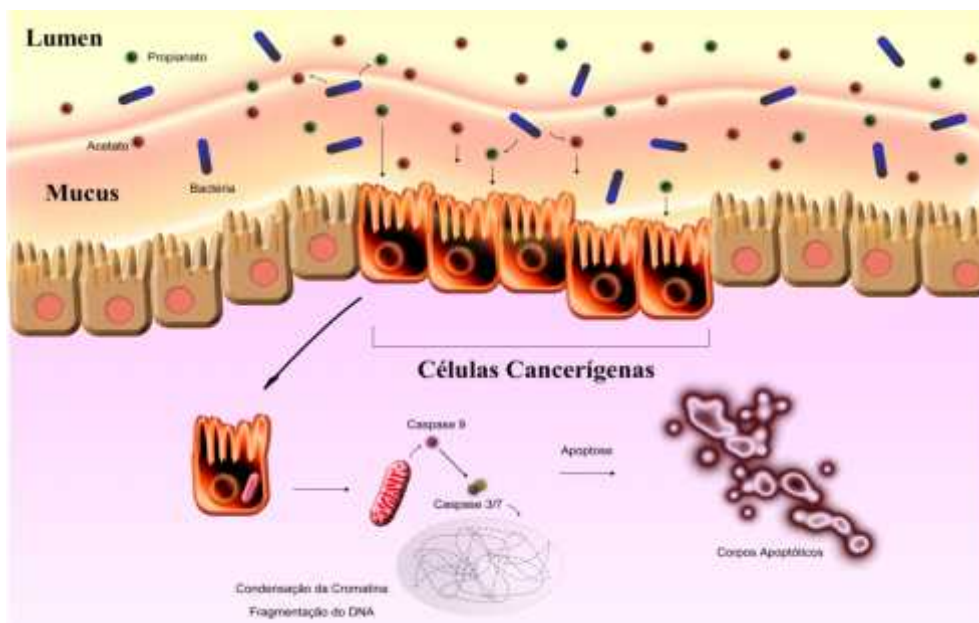


Figura 9: Indução da apoptose a partir de SCFAs produzidos por bactérias propiônicas no cólon.

Além disso, foi proposto que as bactérias propiônicas poderiam complementar o arsenal de opções terapêuticas disponíveis na prevenção e no tratamento de câncer de cólon (Lan *et al.* 2008). Recentemente, esta mesma equipe de pesquisa propôs uma nova formulação de um leite fermentado com algumas linhagens de *P. freudenreichii*, visando uma dieta preventiva ou um suplemento na alimentação de pessoas que apresentam pré-disposição a este tipo de câncer (Cousin *et al.* 2012).

#### **2.4.5. Molécula ácido 1,4-di-hidroxi-2-naftóico (DHNA)**

A molécula ácido 1,4-di-hidroxi-2-naftóico (DHNA) apresenta um acentuado efeito bifidogênico, uma vez que favorece o crescimento de bifidobactérias, propriedade importante para um prebiótico. Os prebióticos são hidratos de carbono de cadeia curta não-digeríveis por enzimas digestivas em seres humanos. Assim, o crescimento de *P. freudenreichii*, que produz DHNA, influencia diretamente no crescimento de bifidobactérias, fazendo de *P. freudenreichii* uma bactéria que possui atividades pro e pre-bióticas. Esta bactéria contribui também para o equilíbrio da flora intestinal, favorecendo o crescimento de bactérias benéficas. De outra maneira, o DHNA permite reduzir a inflamação no cólon (Quigley, Hudson and Englyst 1999; Okada *et al.* 2006). *In vivo*, a administração de DHNA provoca alterações do sistema imune intestinal de camundongos, modificando a resposta imune inata em células intestinais (Okada *et al.* 2013).

Assim, DHNA é uma molécula promissora, uma vez que pode ser utilizada no tratamento de colite. Em um modelo de colite em camundongos, sua administração provoca uma melhoria na histologia do epitélio intestinal, uma supressão da migração de linfócitos e da síntese de interleucinas pró-inflamatórias (Okada *et al.* 2013). Neste contexto, pesquisas suplementares devem ser realizadas objetivando uma otimização de DHNA para aplicações terapêuticas e para demonstrar que as quantidades produzidas *in situ* por *P. freudenreichii* poderiam conduzir aos mesmos efeitos.

#### 2.4.6. Capacidade de adaptação

Esta subseção é destinada a capacidade de sobrevivência de *P. freudenreichii* e sua capacidade de adaptação a diferentes estresses. Esta será dividida em duas partes: tolerância ao estresse térmico e ao estresse digestivo. A tolerância a estes estresses é uma propriedade muito importante em bactéria propiônica para sua manutenção e sua viabilidade, otimizando assim sua utilização como probióticos (Leverrier *et al.* 2003a).

##### Ñ Tolerância ao estresse térmico:

As bactérias propiônicas sobrevivem a uma ampla gama de temperaturas. No caso das etapas da maturação de queijos PPC, estas suportam temperaturas baixas (4°C), médias (30°C) e altas (52-53°). Estas temperaturas podem variar de acordo com as etapas do processamento e o tipo de queijo a ser produzido. Para o queijo Gruyère tradicional, há uma etapa de cozimento que submete as bactérias a uma temperatura bem elevada, entre 52-53°C durante 30 minutos. A termotolerância à 55°C de diferentes linhagens utilizadas na fabricação Gruyère foi estudada através da abordagem proteômica e os perfis obtidos por eletroforese 2D revelaram diferentes respostas de acordo com as linhagens termotolerantes e termosensíveis (Anastasiou *et al.* 2006). Em espécies termotolerantes, proteínas associadas ao processo de fermentação propiônica, metabolismo de aminoácidos, nucleotídeos e fosforilação oxidativa foram superexpressas. Em particular, as chaperonas moleculares e os compostos antioxidantes de tiol são mais induzidas pelo calor. Entretanto, observou-se que para a adaptação térmica as vias de transdução de sinal, biossíntese e manutenção da parede celular contribuíram para a manutenção do processo de fermentação (Anastasiou *et al.* 2006). Outros estudos abordando bactérias propiônicas relataram também que chaperonas moleculares e proteínas de choque térmico são induzidas em muitas situações de estresse (Jan *et al.* 2001; Leverrier *et al.* 2003a). Esta diferença na expressão de proteínas de diversas linhagens de bactérias propiônicas, ressaltam o argumento de que existem fenótipos diferentes dentro da mesma espécie (Anastasiou *et al.* 2006). Esta forma de tolerância é muito importante em *P. freudenreichii* para manter a sobrevivência bacteriana e garantir a fluidez do processo de fabricação de queijos do tipo PPC. Ela consititue também uma vantagem incontestável para sua utilização como probiótico, se associamos sua resistência aos estresses digestivos (Leverrier *et al.* 2003a).

### Ñ Tolerância ao estresse digestivo:

Para que um probiótico exerça sua função é necessário a manter sua atividade metabólica e sobrevivência após a passagem pelo trato gastrointestinal. O trabalho de Lan e colaboradores demonstraram esta capacidade em *P. freudenreichii*, demonstrando a viabilidade destas bactérias no trato digestivo e no cólon de camundongos (Lan *et al.* 2007). Este estudo confirmou também a tolerância ao pH ácido de algumas linhagens específicas de *P. freudenreichii*, como já relatado em estudos anteriores sobre a resistência ao estresse ácido (Jan, Rouault and Maubois 2000) e também a tolerância aos sais biliares (Warminska-Radyko, Laniewska-Moroz and Babuchowski 2002). Experimentos de proteômica, através de eletroforese 2D, revelaram 24 proteínas induzidas durante a adaptação de *P. freudenreichii* em meio contendo sais biliares. Entre elas, estavam proteínas de transdução de sinal e fatores sigma alternativos (Leverrier *et al.* 2003a).

Para avaliar a atividade metabólica de *P. freudenreichii*, uma RT-PCR quantitativa utilizando uma sonda *TaqMan* e selecionando o RNAm codificante para a enzima transcarboxilase foi desenvolvido. Este método não invasivo e eficiente foi utilizado para provar a sobrevivência e a atividade metabólica de *P. freudenreichii* no trato intestinal de camundongos (Hervé *et al.* 2007). Outros métodos buscando avaliar a atividade metabólica de *P. freudenreichii* no queijo foram desenvolvidos no intuito de caracterizar as linhagens quanto à sua sobrevivência. Neste contexto, um método de RT-PCR quantitativa utilizando *SyberGreen* foi otimizado para a quantificação de células e de marcação de RNAm (gênes 16S, *tuf*, *groL*) de *P. freudenreichii* e de *Lactobacillus paracasei*, durante o processo de fabricação de queijos PPC (Falentin *et al.* 2010c). Este método permitiu quantificar a viabilidade celular, bem como instiga estudos para investigar como ocorre a sobrevivência de *P. freudenreichii*.

### 3. Genoma de *Propionibacterium freudenreichii*

Através das tecnologias de sequenciamento de próxima geração (NGS), foi possível a obtenção do sequenciamento de vários genomas de bactérias propiônicas, o que permitiu um melhor conhecimento biológico da espécie. Estes genomas apresentam, em geral, um alto conteúdo GC, em torno de 67% (Falentin *et al.* 2010a) e algumas linhagens podem apresentar plasmídeos em sua composição (Luijk *et al.* 2002). Dados provenientes do sequenciamento podem ser utilizados em estudos comparativos e/ou evolutivos visando compreender a genômica, ao nível estrutural e ao nível funcional deste microorganismo.

Em 2010, o primeiro genoma completo de *P. freudenreichii* foi o da linhagem CIRM-BIA1, o qual foi sequenciado utilizando a metodologia tradicional de Sanger (Falentin *et al.* 2010a). A montagem inicial deste genoma foi realizada sem sequências repetitivas, usando o pacote do software Phred/Phrap/Consed ([www.phrap.org](http://www.phrap.org)) (Vallenet *et al.* 2008; Falentin *et al.* 2010a). Posteriormente, para o fechamento de *gaps*, foi necessário a utilização do *kit GC sequence finishing (GEhealthcare)* e o *kit transposition bombs*, técnicas que permitiram a obtenção do primeiro genoma completo de *P. freudenreichii* pela equipe do STLO - INRA de Rennes, France (Falentin *et al.* 2010a). Este genoma possui um cromossomo circular de 2,7 Mb, alto conteúdo de GC (67%) e encontra-se disponível no banco de dados *Genbank*, com o número de acesso FN806773 (Falentin *et al.* 2010a).

Ao longo dos últimos anos, outros 21 genomas de *P. freudenreichii* foram sequenciados por esta equipe. O sequenciamento destes genomas foi realizado através da plataforma Illumina, utilizando bibliotecas *paired-ends*, com exceção da linhagem CIRM-BIA 9 que utilizou *single reads*. A montagem destes 21 genomas foi obtida através da abordagem *de novo*, e em seguida, os *scaffolds* foram ordenados por comparação ao genoma de referência completo CIRM-BIA1 (Falentin *et al.* 2010a; Loux *et al.* 2015a). Os detalhes da montagem destes 21 genomas são apresentados na Tabela 5. Todos os genomas foram anotados na plataforma AGMIAL, que é uma interface gráfica para anotação de genomas completos e incompletos de procariotos. Os 21 genomas foram depositados no banco de dados do EMBL-EBI e estes dados encontram-se disponíveis em [http://www.ebi.ac.uk/ena/data/view/Taxon:Propionibacterium % 20freudenreichii](http://www.ebi.ac.uk/ena/data/view/Taxon:Propionibacterium%20freudenreichii)) (Loux *et al.* 2015a). Dentre estes 21 genomas, inclui *P. freudenreichii* CIRM-BIA138, uma das linhagens utilizadas durante este trabalho de tese para a análise do RNA-Seq.

Porém, a maioria destes novos genomas de *P. freudenreichii* são ainda incompletos, ou seja, em estado *draft*, isto se deve a dificuldade na etapa de montagem de genomas associada a presença de um elevado número de sequências repetitivas (transpóson).

Análises bioinformáticas foram realizadas em diferentes genomas de *P. freudenreichii*, com o intuito de identificar os genes envolvidos: (i) na resistência a diferentes pH e a presença de sais biliares; (ii) na síntese de vitaminas e aminoácidos e (iii) no armazenamento de carbono. Estas análises demonstram que ela é uma bactéria versátil, adotada de um amplo arsenal de genes para sobreviver em ambientes desfavoráveis (Falentin *et al.* 2010a; Loux *et al.* 2015a). Estes dados contribuíram para a explicação da capacidade de *P. freudenreichii* manter uma atividade metabólica em condições de estresse e de sobreviver a longo prazo. Isto explica em parte sua viabilidade em ambientes que simulam o trato gastrointestinal em humanos e animais. Estas análises genômicas foram também utilizadas para identificar genes potencialmente associados a sua atividade imunomodulatória.

Outros estudos revelaram a presença de mutações pontuais e ilhas genômicas, variando de uma linhagem a outra. Isto demonstra a plasticidade cromossômica desta espécie, que resulta, provavelmente, de transferências horizontais, duplicações, transposições e mutações acumuladas. A análise destas variações entre linhagens, acoplada a uma caracterização fenotípica completa de sua capacidade a metabolizar açúcares permitiu estabelecer uma ligação entre esta plasticidade e o efeito de que determinadas linhagens de *P. freudenreichii* utilizam, preferencialmente, determinadas fontes de carbono. Estas análises abrem também novas perspectivas para o melhoramento de seleção e utilização destas bactérias para a indústria de queijos e de probióticos (Loux *et al.* 2015a).

O genoma completo de *P. freudenreichii* CIRM-BIA1 foi também analisado e comparado ao de *P. acnes*, no intuito de identificar os fatores que diferenciam estas duas espécies próximas. Esta análise também foi realizada para identificar o que estava subjacente as capacidades probióticas de determinadas linhagens de *P. freudenreichii* (Falentin *et al.* 2010a).

As análises genômicas estruturais baseadas no DNA são completadas à análises funcionais através de estudos sobre o RNA, proteínas e produtos do metabolismo desta bactéria, respectivamente, abordagens de transcriptômica, proteômica e metabolômica. Isto será ilustrado no segundo capítulo da seção de resultados deste manuscrito de tese:



utilização da genômica funcional para a compreensão do fenômeno de sobrevivência à longo prazo em *P. freudenreichii*.

**Tabela 5: Genomas depositados de *P. freudenreichii* pela equipe do STLO.**

Number in CIRM-BIA collection (Rennes, France)	Locus_tag	Isolated from (date)	Bioproject	Raw reads accession number	Read quantity	Read length	Sequencing depth	Number of scaffolds >1000 bp	Cumulative contig length	N50	Number of N segments	Number of complete genes
9	PFCIRM9	Emmental cheese (1989)	<b>PRJEB6427</b>	ERS638416	26 013 151	37	370	204	2 533 825	21 887	0	2 308
118	PFCIRM118	Gruyère cheese (1973)	<b>PRJEB6428</b>	ERS638410	51 926 392	37	739	41	2 583 496	193 767	59	2 325
119	PFCIRM119	Gruyère cheese (1973)	<b>PRJEB6430</b>	ERS638409	58 084 132	37	827	46	2 644 402	128 788	91	2 422
121	PFCIRM121	Swiss cheese (1937)	<b>PRJEB6431</b>	ERS638414	52 432 380	37	746	59	2 561 530	86 244	68	2 304
122	PFCIRM122	NCIB (1992)	<b>PRJEB6432</b>	ERS638421	54 950 090	37	782	78	2 614 705	73 741	70	2 348
123	PFCIRM123	Morbier cheese (1992)	<b>PRJEB6438</b>	ERS638423	54 523 668	37	776	47	2 604 827	114 519	78	2 377
125	PFCIRM125	Emmental cheese (1992)	<b>PRJEB6434</b>	ERS638418	54 868 488	37	781	43	2 511 448	238 430	59	2 272
127	PFCIRM127	Emmental cheese (1992)	<b>PRJEB6435</b>	ERS638428	55 178 920	37	785	57	2 567 344	81 280	53	2 330
129	PFCIRM129	Emmental cheese (1992)	<b>PRJEB4826</b>	ERS638425	52 269 870	37	744	59	2 591 314	123 180	59	2 338
134	PFCIRM134	NA	<b>PRJEB6441</b>	ERS638419	50 523 582	37	719	47	2 635 396	153 798	152	2 318
135	PFCIRM135	ewe raw milk (1994)	<b>PRJEB6442</b>	ERS638422	50 649 408	37	721	47	2 620 043	114 923	75	2 343
138	PFCIRM138	Emmental cheese (1992)	<b>PRJEB6433</b>	ERS638424	58 100 372	37	827	72	2 573 580	80 515	110	2 304
139	PFCIRM139	Emmental cheese (1992)	<b>PRJEB6436</b>	ERS638426	60 700 312	37	864	45	2 589 117	106 963	53	2 362
456	PFCIRM456	raw milk Raclette cheese (1992)	<b>PRJEB6445</b>	ERS638427	49 094 398	37	699	38	2 502 068	131 281	61	2 253
508	PFCIRM508	Gruyère (1973)	<b>PRJEB6429</b>	ERS638411	50 259 834	37	715	47	2 578 004	141 315	93	2 302
512	PFCIRM512	raw milk Morbier cheese (1994)	<b>PRJEB6439</b>	ERS638413	63 808 750	37	908	45	2 580 419	153 592	57	2 349
513	PFCIRM513	egyptian ras cheese (1995)	<b>PRJEB6440</b>	ERS638415	62 230 134	37	886	48	2 659 761	179 407	54	2 416
514	PFCIRM514	hay (1994)	<b>PRJEB6443</b>	ERS638417	51 481 488	37	733	51	2 521 626	121 592	66	2 285
516	PFCIRM516	nepalian yack cheese (2000)	<b>PRJEB6444</b>	ERS638420	61 823 266	37	880	62	2 630 828	125 877	205	2 342
527	PFCIRM527	Fribourg cheese (1992)	<b>PRJEB6437</b>	ERS638412	69 806 362	37	993	63	2 576 369	104 864	57	2 344
1025	PFCIRM1025	Emmental cheese (1992)	<b>PRJEB6446</b>	ERR738371	61 323 626	81	1 910	104	2 699 685	62 523	19	2 308

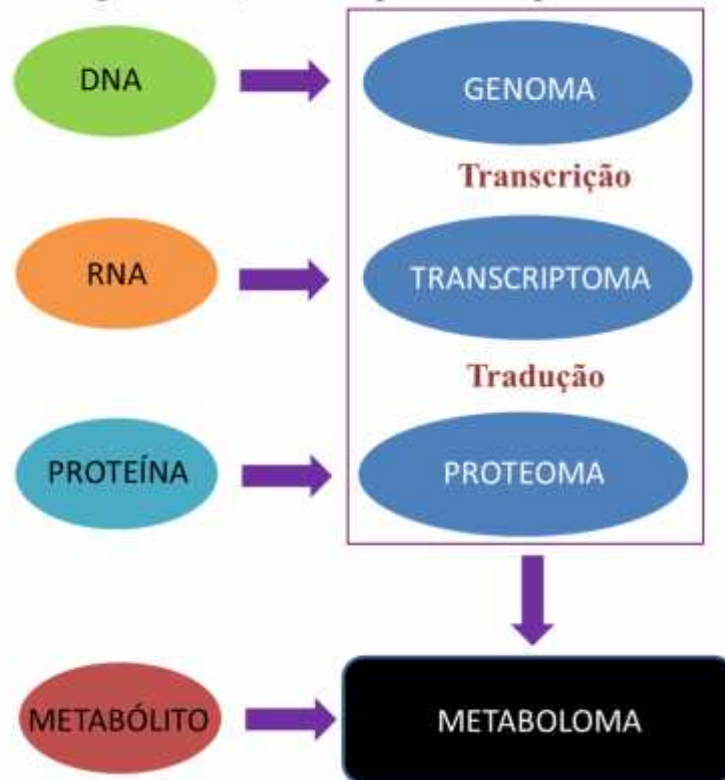
#### 4. Transcriptômica de *Propionibacterium freudenreichii*

Nos últimos 10 anos, vários estudos abordaram a expressão gênica em *Propionibacterium*, no intuito de melhor compreender e de explicar como ocorre a regulação destes genes neste grupo bacteriano. Estes dados transcriptômicos são depositados em bancos de dados, como o EMBL-EBI (<http://www.ebi.ac.uk/genomes/>), criados para permitir o acesso público, fornecendo dados de genoma e transcriptoma de diferentes bactérias. Estes trabalhos nos auxiliam a compreender como as bactérias propiônicas se adaptam e regulam a expressão de genes em resposta a diferentes condições ambientais.

Espécies como *P. acnes* e *P. freudenreichii* são de grande importância neste gênero, uma vez que *P. acnes* é comensal e faz parte da microbiota da pele; enquanto que *P. freudenreichii*, como falado anteriormente, é de interesse industrial de laticínios no setor de queijos PPC e de interesse biotecnológico para a síntese de biomoléculas e por suas propriedades probióticas. Os primeiros estudos de *P. freudenreichii* foram realizados com as técnicas de microarranjo e RT-qPCR. Estes estudos demonstraram a sobrevivência destas bactérias à temperaturas variadas durante o processo de maturação de queijos PPC, assim como sua adaptação metabólica (*in situ*) no cólon de porcos (Dalmasso *et al.* 2012a; Saraoui *et al.* 2013a). A fim de ampliar os conhecimentos a cerca da sobrevivência desta bactéria, utilizamos o RNA-Seq, uma abordagem transcriptômica mais completa, associada à uma análise metabolômica direcionada. Nós apresentaremos na seção II de resultados deste manuscrito, um artigo a respeito da análise de *P. freudenreichii* CIRM-BIA 138, integrando seus dados transcriptômicos e metabolômicos obtidos *in vitro*.

A associação destes dados a níveis de RNA e bioquímica, aos dados proteômicos já obtidos de *P. freudenreichii*, permitirá de abordar o estudo funcional desta bactéria e nos esclarecerá estas capacidades de adaptação e de sobrevivência (Figura 10).

**Correlação dos dados genômicos, transcriptômicos e proteômicos de *P. freudenreichii***



**Figura 10:** Esquema de correlação dos dados de genoma, transcriptoma e proteoma para a compreensão do metaboloma desta bactéria.

## 5. Principais vias metabólicas de *Propionibacterium freudenreichii*

As bactérias propiônicas obtêm sua energia através de diversos substratos de carbono através da fermentação propiônica. Em *P. freudenreichii*, a associação dos ciclos de Krebs e de Wood-Werkman são características da espécie.

O ciclo de Krebs, denominado também de ciclo do ácido cítrico ou de ciclo dos ácidos tricarboxílicos (TCA - termo em inglês), corresponde a reações metabólicas que ocorrem no citoplasma da célula bacteriana. Este ciclo é uma via anfipática apresentando etapas de (i) degradação dos metabólitos para a geração de energia –ATP– (via catabólica), e (ii) produção de precursores para a biossíntese (via anabólica) (Nelson *et al.* 2008).

Em *P. freudenreichii*, o ciclo de Wood-Werkman, sendo conhecido também como via do carbono central, é um componente central do metabolismo. Este ciclo permite a formação de propionato à partir do piruvato (Wood 1981). Durante a sua fermentação, as bactérias propiônicas são capazes de utilizar o lactato e de produzir, como citado anteriormente, determinados subprodutos de interesse para a composição do Emmental. A figura 11 ilustra o metabolismo de *P. freudenreichii*, baseado no ciclo de Wood-Werkman e sua associação ao ciclo de Krebs.

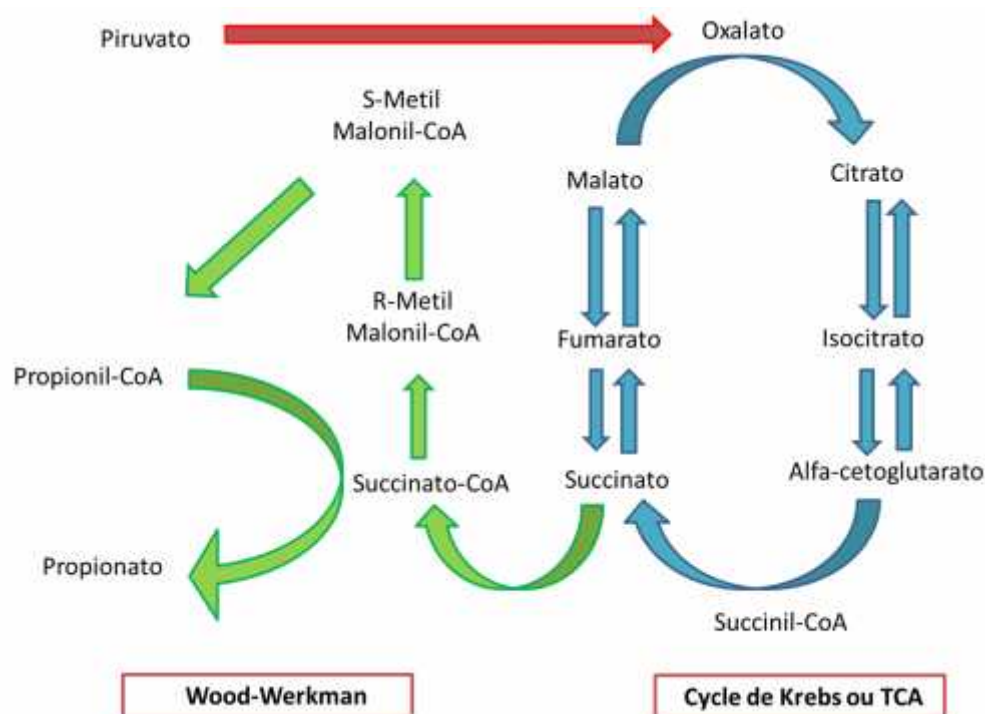
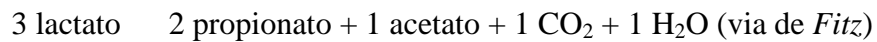


Figura 11 : Via metabólica de Wood-Werkman em *P. freudenreichii* acoplada ao ciclo de Krebs. Figura adaptada do artigo: A Temporal -omic Study of *Propionibacterium freudenreichii* CIRM-BIA1<sup>T</sup> Adaptation Strategies in Conditions Mimicking Cheese Ripening in the Cold (disponível em <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0029083>).

A reação global descrita nestas bactérias segue o seguinte equilíbrio: (Crow 1986)



*P. freudenreichii* tem um papel fundamental na maturação de queijos PPC e no desenvolvimento de seus sabores (Figura 12). Também as diversas vias metabólicas resultam na formação de compostos aromáticos, os quais foram pesquisados, especialmente, combinando estudos genômicos e metabolômicos. Diversas vias metabólicas, como o metabolismo da glicólise, da lipólise, dos aminoácidos e do armazenamento de energia foram reconstruídos integralmente para determinadas linhagens. Estes trabalhos foram alvos de muitos artigos e revisões detalhadas (Thierry *et al.* 2011).



**Figura 12:** Esquema de processos metabólicos envolvidos na maturação dos queijos PPC visando a obtenção de compostos aromáticos.

**INTRODUCTION  
BIBLIOGRAPHIQUE**

Cette partie du manuscrit est consacrée à la présentation de données générales sur le genre *Propionibacterium* et à celle, plus détaillée, de l'espèce *P. freudenreichii*, concernant son application dans l'industrie, ses propriétés métaboliques et probiotiques, et sa capacité à résister aux différents stress.

## **1. Genre *Propionibacterium***

### **1.1. Classification, écologie et physiologie**

En 1906, Von Freudenreichi et Orla-Jensen ont rapporté le premier isolement de bactéries propioniques à partir de fromage Emmental (von Freudenreich and Orla-Jensen 1906). En 1909, le nom du genre *Propionibacterium* a été proposé par Orla-Jensen (Cummins and Johnson, 1986), en raison de la capacité de cette bactérie à produire de l'acide propionique pendant le processus de fermentation (Orla-Jensen 1909). Plus tard, en 1946, Douglas et Gunter ont proposé l'inclusion de bactéries appartenant à l'espèce *Corynebacterium acnes* au sein du genre *Propionibacterium*. Et, par la suite, en 1963, d'autres auteurs ont confirmé cette reclassification, faisant de *C. acnes* une espèce du genre *Propionibacterium* (Moore and Cato 1963).

Les bactéries du genre *Propionibacterium* sont des bactéries à Gram positif, appartenant à l'ordre des Actinomycétales, dans le phylum Actinomycètes. Elles sont immobiles et ne forment pas de spores. Ce genre peut être divisé en deux grands groupes selon son habitat : (i) les produits laitiers, dans lesquels on trouve les bactéries propioniques lactières ; (ii) la peau et/ou les muqueuses, sur laquelle on trouve les bactéries cutanées ou commensales. Le groupe des bactéries propioniques lactières, comprend différentes espèces, dont les plus connues sont *P. freudenreichii*, *P. thoenii*, *P. jensenii* et *P. acidipropionici*. Ces espèces ne sont pas pathogènes pour l'homme et sont généralement isolées des produits fermentés alimentaires, comme l'ensilage, la fermentation d'olives, les olives, la distillerie de rhum. Quelles études rapportent leur présence dans des échantillons de l'environnement, par exemple, des sols, et des animaux notamment dans le tractus intestinal des rats (Cummins and Johnson 1986; Mantere-Alhonen 1982; Balows 1992). D'autres espèces telles que *P. cyclohexanicum* (Kusano *et al.* 1997) et *P. microaerophilum* (Koussémon *et al.* 2001) ont été plus récemment décrites et classées dans des groupe des bactéries propioniques lactières non-pathogènes, bien qu'ayant été isolées non pas de produits laitiers, mais d'aliments



avariés ou de sous-produits. Des exemples de ces occurrences sont *P. cyclohexanicum*, trouvé dans le jus d'orange fermenté et *P. microaerophilum*, décrit dans les eaux usées provenant de pressoirs à olives. Toutefois, dans le groupe de *Propionibacterium* commensal ou cutané, il y a trois espèces, dont certaines comme *P. acnes*, *P. avidum* et *P. propionicum*, provoquent des pathologies chez l'homme. Les espèces *P. lymphophilum* et *P. granulosum* appartiennent aussi à la classification des commensales, néanmoins elles représentent un groupe distinct, étant phylogénétiquement loin des autres groupes (Meile *et al.* 1999).

## 1.2. La caractérisation et l'importance du micro-organisme étudié : *Propionibacterium freudenreichii*

Les conditions optimales de laboratoire pour la croissance de *P. freudenreichii* sont le milieu YEL (bouillon lactate et extrait de levure), proposé par Malik et collaborateurs (Malik, Reinbold and Vedamuthu 1968), et à une température d'incubation de 30°C. Il s'agit d'une bactérie anaérobie facultative qui peut être trouvée *in vitro* sous diverses formes (pléomorphe) : cocciforme, bifide ou ramifiée (Figure 1).

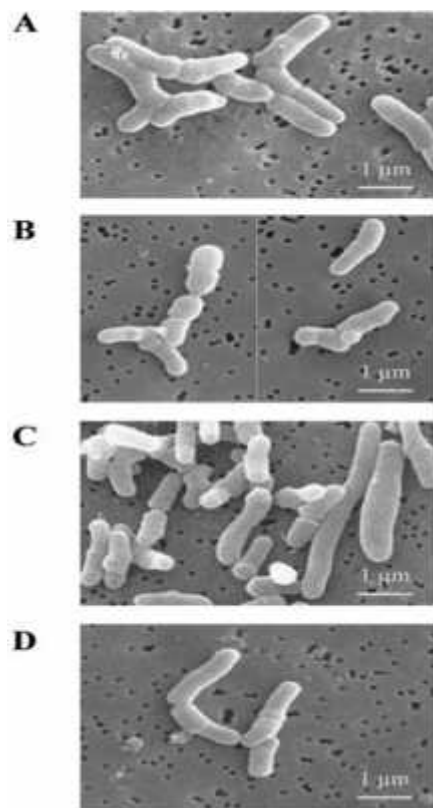


Figure 1 : Aspect pléomorphe de *P. freudenreichii* pendant le stress acide : (A) cellules non traitées; (B) cellules soumises à un pH = 2,0; (C) cellules soumises à un pH = 5,0; (D) cellules soumises à un pH = 5,0 et soumis ensuite à un pH = 2,0, visualisées par le microscopie électronique à balayage (Jan *et al.*, 2001).

Auparavant, l'espèce *P. freudenreichii* était divisée en deux sous-espèces : *P. freudenreichii* subsp. *freudenreichii*, qui est incapable de fermenter le lactose, mais peut réduire le nitrate ; et *P. freudenreichii* subsp. *shermanii*, qui fermente le lactose sans réduire le nitrate (Cummins and Johnson, 1986). Cependant, des études récentes montrent qu'il existe des souches de *P. freudenreichii* qui présentent les deux phénotypes : le lactose et le nitrate réductase positif ou le lactose et le nitrate réductase négative (de Carvalho, Gautier and Grimont 1994; Vorobjeva 1999; Dalmaso *et al.* 2011). En outre, dans le génome complet de *P. freudenreichii* CIRM-BIA 1, il a été observé que le gène *lacZ*, responsable de l'activité beta-galactosidase, est situé dans une région de transposase et qu'il y a des pseudogènes dans la région du génome qui code l'information nécessaire à la réduction du nitrate (Falentin *et al.* 2010a). Actuellement, il n'y a pas de classification des sous-espèces de *P. freudenreichii* au niveau phénotypique et d'autres recherches pour explorer le génome de cette bactérie doivent encore être effectuées. Ceci est d'une grande importance dans l'industrie alimentaire, puisque que ces propriétés biochimiques peuvent modifier la qualité finale d'un produit, en l'occurrence les fromages PPC.

*P. freudenreichii* présente une parfaite innocuité sanitaire et possède deux statuts importants pour son utilisation : le *Generally Recognized As Safe* (GRAS), délivré par l'Agence américaine des produits alimentaires et médicaux (Food and Drug Administration - FDA) (Mogensen *et al.* 2002) ; et la *Qualified Presumption of Safety* (QPS), délivré par l'Autorité européenne de sécurité des aliments (European Food Safety Authority - EFSA) (EFSA EXPERTS, 2008). Ces statuts sont approuvés par les autorités du secteur visant à contrôler la qualité des produits, ce qui garantit la sécurité et autorise l'utilisation de ces bactéries dans l'industrie alimentaire, par exemple, dans la fabrication du fromage Emmental. Ces bactéries sont très largement répandues dans la nature, on les trouve dans le tractus gastro-intestinal si elles ont été ingérées, et elles sont aussi présentes dans de nombreux aliments fermentés, du fait de leurs caractéristiques qui les rendent résistantes au pH acide (Jan, Rouault and Maubois 2000; Huang and Adams 2004; Suomalainen *et al.* 2006), et aux basses températures dans le processus d'affinage des fromages PPC (Dalmaso *et al.* 2012a). La concentration généralement rencontrée dans l'Emmental pour ces bactéries est de l'ordre de  $10^9$  UFC/g. Ces bactéries possèdent des qualités très importantes pour diverses applications dans le domaine industriel, qui seront présentées dans la section suivante.

## 2. Applications industrielles de *Propionibacterium freudenreichii*

Les aliments fermentés, en particulier les produits laitiers fermentés, sont consommés par l'homme depuis plusieurs milliers d'années. Depuis le Moyen Age, la fabrication des fromages de type Emmental dans les alpages de Savoie est décrite dans des manuscrits. Au XIX<sup>e</sup> siècle, apparaissent les premiers rapports sur les "fruitières" qui sont des lieux d'exploitation et de transformation du lait en fromage, regroupant des producteurs mettant en commun le « fruit » de leur production d'été. Puis, au XX<sup>e</sup> siècle, une forte industrialisation a été observée, conduisant à une délocalisation de la production de ces fromages.

A l'instar de l'Emmental, tous les produits laitiers qui auparavant étaient produits de manière plutôt empirique ont bénéficié des progrès de la microbiologie et de la technologie alimentaire, ce qui a permis de mieux maîtriser, diversifier et sécuriser leur production.

Les principales espèces de bactéries utilisées aujourd'hui pour la production de produits laitiers fermentés appartiennent aux genres : *Bifidobacterium*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Streptococcus* et *Propionibacterium* (Tableaux 1 et 2). Ces genres jouent des rôles importants dans l'industrie laitière, aussi bien comme dans la synthèse de biomolécules, que comme dans le développement de produits probiotiques.

Dans ce contexte, *P. freudenreichii* joue un rôle important dans la fabrication et particulièrement dans l'affinage des fromages PPC. Cette espèce est aussi utilisée pour sa capacité à produire des bactériocines, des composés de l'acide linoléique, de l'acide propionique, du propène, du tréhalose et des vitamines du groupe B, comme par exemple, la vitamine B12 (Piao *et al.* 2004; Poonam *et al.* 2012). Certaines souches possèdent également des propriétés probiotiques et peuvent moduler la composition du microbiote intestinal en produisant des facteurs bifidogènes, ou moduler le niveau d'inflammation dans le contexte de maladies inflammatoires chroniques de l'intestin (MICI). Elles sont d'autant plus intéressantes qu'elles sont capables de résister au stress du tractus gastro-intestinal.

Comme décrit précédemment, cette bactérie présente la capacité de s'adapter à des stress de nature variable, tels que le pH, la température et les sels biliaires. Cette robustesse confère un avantage à cette espèce qui lui permet de rester viable et de maintenir des activités métaboliques dans divers environnements. Ces caractéristiques sont importantes pour ses applications dans l'industrie alimentaire, tout particulièrement dans l'industrie fromagère des PPC ; en technologie, pour la synthèse de

biomolécules et/ou pour l'aliment-santé (probiotiques), toutes ces applications seront présentées en détail dans cette section.

**Tableau 1 : L'action microbienne dans les aliments laitiers fermentés (tableau adapté <http://mastervrv.free.fr/S3/BA/Cours/BA-FERMENTATION.pdf>).**

Effets sur le produit	Exemples	Microorganismes impliqués
Modification de la texture	Coagulation du lait	Plusieurs bactéries lactiques
	Texture filante ou veloutée de produits laitiers frais fermentés	Bactéries lactiques (dextranes)
	Formation des «trous» dans le fromage	Bactéries propioniques
Aromatisation	Goût typique du beurre (diacétyl)	<i>Lactococcus lactis diacetylactis</i>
	Goût typique du yaourt (acétaldéhyde et acide lactique)	<i>Lactobacillus delbrueckii bulgaricus</i>
	Développement du goût du cacao	Plusieurs bactéries
Coloration	Brunissement de la pâte des fromages PPC* par le propionate de calcium	Bactéries propioniques
Modification du goût	Acidification du chou (choucroute)	<i>Lactobacillus</i>
	Désacidification du vin	Bactéries malolactiques
Stabilisation microbienne	Acidification	Bactéries lactiques et acétiques

\*PPC = pâte pressée cuite

**Tableau 2 : Principaux ferments pour l'industrie laitière et fromagère (tableau adapté <http://mastervrv.free.fr/S3/BA/Cours/BA-FERMENTATION.pdf>).**

Ferments	Applications
<i>Lactobacillus bulgaricus</i> <i>Streptococcus thermophilus</i> <i>Bifidobacterium sp.</i>	Yaourts
<i>Lactobacillus acidophilus</i>	Laits fermentés
<i>Lactobacillus caucasius</i>	Kéfir (boisson fermentée)
<i>Streptococcus cremosnis</i> <i>Streptococcus lactis</i> <i>Streptococcus diacetylactis</i> <i>Leuconostoc cremonis</i> <i>Lactobacillus cremonis</i> <i>Lactobacillus lactis</i>	Premières étapes (caillage) de la plupart des fromages Fromages non affinés à pâte fraîche Beurre
<i>Propionibacterium shermanii</i> <i>Propionibacterium freudenreichii</i>	L'affinage des fromages PPC*

\*PPC = pâte pressée cuite

Dans les paragraphes à suivre, nous présentons quelques données sur l'importance de *P. freudenreichii* dans les applications industrielles :

- **Fromages PPC** → les bactéries propioniques sont nécessaires pour l'affinage des fromages PPC. Leur survie est nécessaire pour assurer la saveur et la typicité de ces fromages.
- **Biomolécules** → bactéries propioniques permettent la production de diverses biomolécules d'intérêt industriel, comme par exemple, l'acide propionique et la vitamine B12, produits du métabolisme de ces bactéries.
- **Probiotiques** → le maintien de l'activité métabolique de *P. freudenreichii* dans les produits probiotiques est d'une grande importance, car cette bactérie peut agir sur le tractus gastro-intestinal via, par exemple, l'immunomodulation. Pour cela, il est nécessaire que les bactéries survivent pendant de longues périodes en passant par des étapes de transformation et de stockage, sans altérations majeures sur la viabilité et l'activité visée.

## 2.1. Industrie des fromages PPC (pâte pressée cuite)

Les fromages sont produits partout dans le monde et sont classés en fonction de leurs caractéristiques, tels que la consistance de la pâte, la teneur en matière grasse, le type de lait utilisé, le type de coagulation, le type de croûte et le temps d'affinage. *P. freudenreichii* est utilisé conjointement avec d'autres cultures de bactéries, principalement dans la fabrication du fromage Emmental. Ce groupe est le troisième segment le plus important dans l'industrie du fromage, et est appelé « pâte pressée cuite » (PPC), il comporte d'autres fromages, comme le Comté. Le terme PPC correspond aux fromages dont la coagulation et la déshydratation sont prolongées et qui sont ensuite pressés au moment du moulage pour éliminer la quantité maximum de lactosérum, on les appelle également "fromages à pâte dure". Après l'étape de coagulation et de tranchage, le caillé est « cuit » par une étape de chauffage au-dessus de 50°C pendant 30 à 60 minutes dans le sérum, ce qui permet d'obtenir, après égouttage et pressage, un extrait sec élevé et une plus longue conservation.

Selon le Centre National Interprofessionnel de l'Economie Laitière en 2014, la France a produit plus de 260.000 tonnes d'Emmental. La région Grand Ouest de la France, comprenant la Bretagne, le Pays de Loire et la Basse-Normandie, représente les deux tiers de cette production (CNIEL, 2014). La Bretagne est la première région productrice de lait de vache et d'Emmental en France (Figure 2).

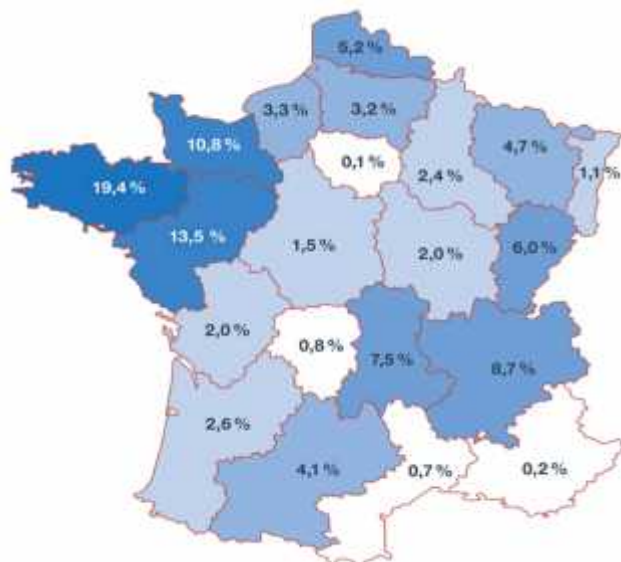


Figure 1 : La production du lait de vache sur le territoire français (La géographie du lait en 2013 - Enquête Annuelle Laitière).

Ces produits sont très importants pour l'économie laitière de la France, particulièrement l'Emmental, qui est le grand représentant de ce secteur (Tableau 3).

**Tableau 3: La production des fromages PPC en tonnes (L'Enquête Annuelle Latière).**

<b>Fromages PPC</b>	<b>2011</b>	<b>2012</b>	<b>2013</b>
Emmental	256.980	253.225	252.892
Comté AOP	58.114	57.561	59.459
Beaufort AOP	5.044	4.919	4.787
Gruyère IGP	-	2.679	2.929
Autres PPC	-	8.913	7.938
Total PPC	333.258	327.298	328.006

Pour comprendre le processus de fabrication de ces fromages PPC, le tableau 4 résume quelques étapes importantes, et la figure 3 illustre l'ensemble du processus.

**Tableau 4: Les étapes de la fabrication des fromages.**

Étapes	Description
<b>La pasteurisation</b>	Le chauffage du lait à des températures élevées. Cette étape vise à éliminer les microorganismes indésirables. Cependant, cette étape n'est pas nécessaire pour les fromages fabriqués avec du lait cru ou pasteurisé.
<b>La coagulation</b>	Cette étape vise à la concentration de la protéine du lait et peut être divisée en deux sous-étapes: (i) acide et (ii) présure. La première est obtenue à partir de la production d'acide lactique par les bactéries ou par l'addition d'acides organiques dans le lait. La deuxième, l'étape enzymatique, se produit à partir de l'activité de la flore microbienne du lait ou par l'ajout de l'extrait de présure.
<b>Le moulage</b>	Les caillés obtenus à l'étape précédente sont placés dans différents moules, en fonction du fromage désiré.
<b>L'égouttage</b>	Le but de cette étape est de séparer le lactosérum du lait caillé, ce qui prolonge la conservation du fromage. Certains fromages peuvent être consommés après ce processus.
<b>Le salage</b>	Les fromages frais sont démoulés et salés avec du sel ou de la saumure saturée en sel. Cela assure une action antiseptique et une meilleure conservation et renforce la saveur au fromage.
<b>L'affinage</b>	Cette étape peut prendre des jours ou des mois, selon la variété de fromage. Ceux-ci seront stockés et recevront des soins en cave, où la température et l'humidité sont contrôlées.



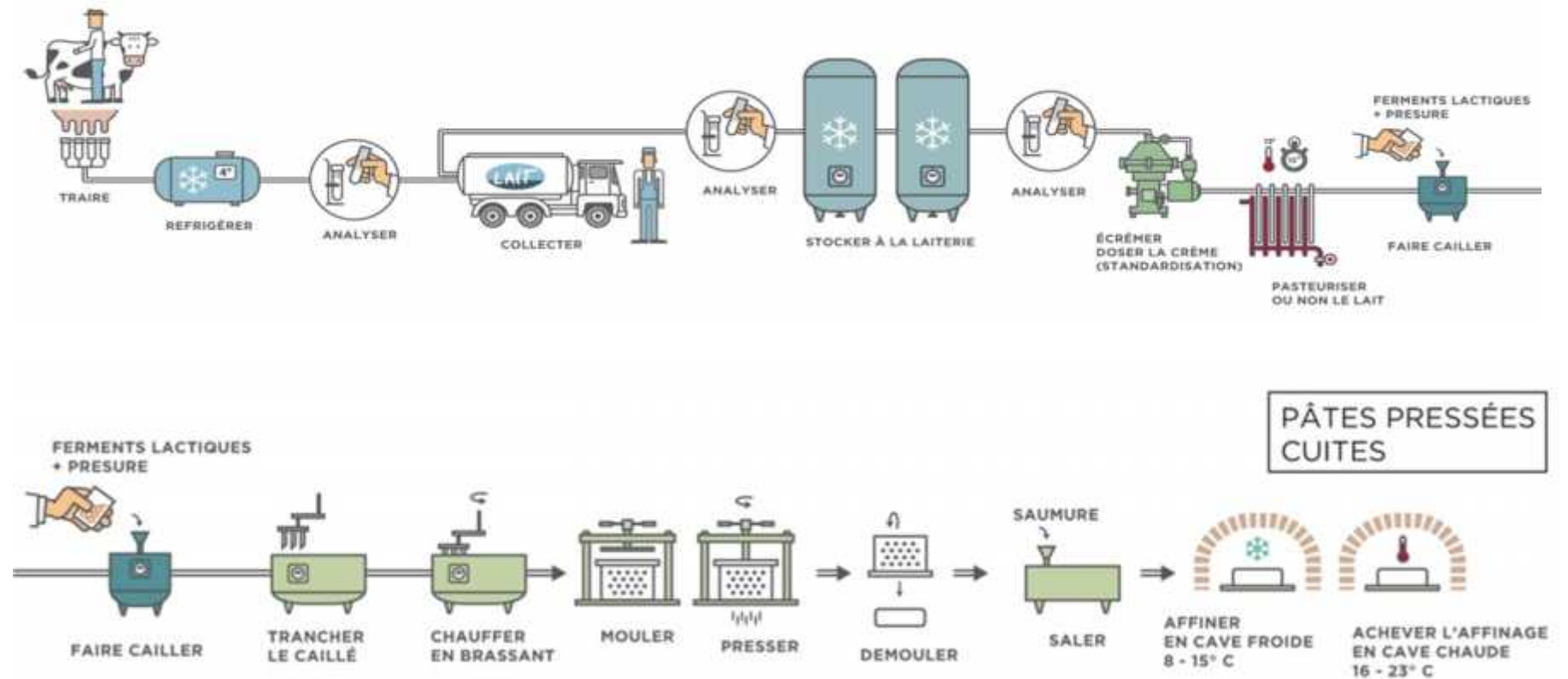


Figure 3 : Processus de transformation des fromages PPC (<http://www.produits-laitiers.com/>). Le processus de fabrication des fromages PPC commence par la collecte du lait dans les machines automatiques des exploitations laitières. Ensuite, ce lait est refroidi et analysé conformément à la réglementation pour la sécurité des consommateurs, afin d'assurer un produit final de bonne qualité. Ce lait est transporté dans des camions citernes isothermes à l'industrie laitière où commencent toutes les étapes de traitement et de fabrication du fromage.

L'Emmental est un fromage de poids dont les meules pèsent entre 60 et 120 kg. Il faut 12 litres de lait pour obtenir 1 kg de ce fromage (CNIEL, 2015). Pendant son affinage de nombreux paramètres varient : la température, le pH et l'activité de l'eau. *P. freudenreichii* joue un rôle important dans son affinage (Figure 4), produisant par fermentation de l'acétate et du propionate, qui participent la flaveur de l'Emmental, et du CO<sub>2</sub>, responsable de l'ouverture du fromage par la formation de trous dans la pâte (les « yeux ») (Langsrud and Reinbold 1973; Meurice *et al.* 2004).

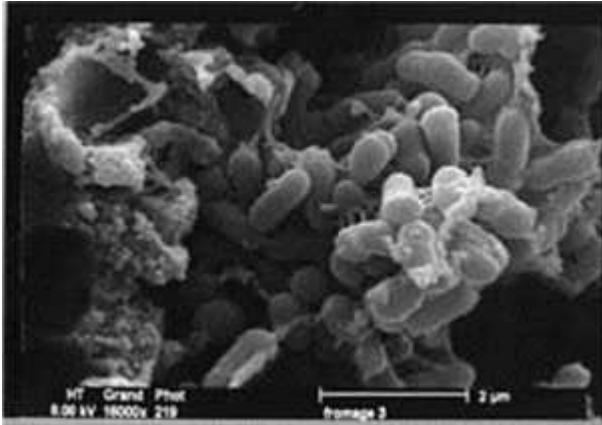


Figure 4 : *P. freudenreichii* dans le fromage Emmental (© INRA).

*P. freudenreichii* joue ainsi un rôle important dans la composition de l'arôme et de la saveur de certains produits laitiers, parmi lesquels l'Emmental. En effet, pendant l'affinage prolongé de ce fromage, *P. freudenreichii* assure des réactions de lipolyse. La lipolyse contribue à la formation de saveurs et est réalisée par des estérases qui sont capables d'effectuer l'hydrolyse des triacylglycérides du lait en libérant des acides gras à chaîne courte, à chaîne moyenne et à chaîne longue (Figure 5). Ceux sont eux qui donnent aux fromages leurs différentes sensations gustatives, comme par exemple, une saveur forte, épicée et/ou salée (Berdagué *et al.* 1990).

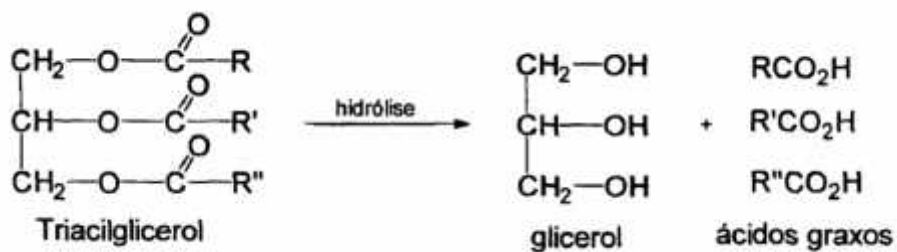


Figure 5 : Schéma de réaction d'hydrolyse de triacylglycérides.

Les études de Dherbécourt et collaborateurs ont identifié 23 estérases (estérases putatives) chez *P. freudenreichii*, potentiellement impliquées dans ce processus d'affinage.

Cette identification s'est faite à partir de la combinaison de méthodes classiques et de l'étude génomique de cette bactérie (Dherbécourt *et al.* 2008). Pour l'approche génomique, les outils bioinformatiques et les méthodes utilisés ont été : (i) la recherche automatique d'estérases par homologie, (ii) la recherche des motifs protéiques dans l'InterProScan, (iii) la recherche manuelle par homologie et (iv) la recherche manuelle faite par motifs spécifiques grâce à Prosite. Ces méthodes ont été indispensables pour une première sélection d'enzymes d'intérêt. Ainsi 12 gènes candidats codant pour des estérases potentielles présentant des motifs GX SXG caractéristiques des / hydrolases ont été sélectionnés et clonés dans *E. coli*. Des zymographes ont également été effectués afin de caractériser l'activité de ces estérases. L'inactivation et la complémentation des deux gènes d'estérases codant pour des protéines ancrées en surface de *P. freudenreichii* ont révélé que le gène pf279 était responsable de la lipolyse. De plus, l'analyse du génome de 21 souches de *P. freudenreichii* a montré que ce gène pouvait présenter des mutations non-sens dans quelques souches qui ne présentent pas d'activité lipolytique (Abeijon Mukdsi *et al.* 2014).

Nous avons déjà vu que *P. freudenreichii* survivait dans divers environnements et contraintes de façon souche dépendante. L'importance de ce phénomène donne de l'attrait à l'étude des stratégies qui permettent la survie et le maintien d'une activité métabolique pendant de longues périodes. Afin d'exploiter ces mécanismes de survie à long terme, les études qui associent des expériences microbiologiques (*in vitro*) en lien avec des applications industrielles et biotechnologiques sont d'une grande importance. Ces études seront décrites dans les sections suivantes.

## 2.2. Optimisation de la production de vitamine B12

En 1911, le terme « vitamine » a été proposé par Casimir Funk pour désigner les amines qui sont essentiels à la vie (vita-amine). Plus tard, cette nomenclature a commencé à être utilisée également pour les substances organiques essentielles, sans valeur énergétique mais actives en petites quantités, contribuant à l'entretien du métabolisme d'un organisme (Florent 1986; Vandamme 1992).

La vitamine B12, aussi appelée cobalamine ou cyanocobalamine, a été isolée pour la première fois en 1948 mais elle n'a été synthétisée qu'en 1973. Elle est en effet difficile à obtenir et sa synthèse comporte plusieurs étapes pour obtenir la molécule finale (Vandamme 1989). Comme cette vitamine est impliquée dans des processus essentiels pour la régulation du métabolisme des animaux et des humains comme la formation et maturation des érythrocytes et la maintenance du système nerveux, les stratégies alternatives visant à l'optimisation de sa production sont d'un grand intérêt.

Dans ce contexte, l'obtention de la vitamine B12 en utilisant des micro-organismes est devenue une solution prometteuse pour le marché industriel. Parmi les bactéries qui en produisent, *P. freudenreichii* se distingue en raison de son bon rendement. Ainsi, une concentration de 5,12 mg/L peut être obtenue dans le fromage après 168h de fabrication (Marwaha, Sethi and Kennedy 1983). Des études d'amélioration de la production de vitamine B12 chez *P. freudenreichii* ont montré qu'il est possible de parvenir à un rendement allant jusqu'à 6,35 mg/L après 192h de fermentation (García Hernández and Pérez Mendoza 1983).

De nouvelles approches ont donc été proposées pour la biosynthèse de la vitamine B12 au cours des décennies. L'obtention de la vitamine B12 par fermentation de *P. freudenreichii* a été analysée dans le travail de Piao et collaborateurs, qui ont montré qu'en jouant sur l'expression des gènes *cobA*, *cbiLF* et *cbiEGH*, impliqués dans la synthèse de la vitamine B12, on pouvait augmenter la production de vitamine B12 (Piao *et al.* 2004). D'autres études ayant le même but, comme le travail de Murooka et collaborateurs, ont exprimé et amplifié les gènes de l'opéron *cob* et *cbi*, pour augmenter la biosynthèse des cobalamines pendant le processus de fermentation (Murooka *et al.* 2005).

Pour optimiser la production de vitamine B12, une autre étude a été réalisée par Wang et collaborateurs se concentrant sur l'effet de l'acide propionique et du 5,6-diméthylbenzimidazole au cours de la biosynthèse de cette vitamine au cours de la culture de *P. freudenreichii* (Wang *et al.* 2015). Il a été montré que la concentration d'acide propionique varie de la phase initiale (10 à 20 g/L) à la phase tardive (20 à 30 g/L). Ceci

conduit à une rétroaction négative sur la biosynthèse de la vitamine B12, mais peut être modifié par l'addition de 5,6-diméthylbenzimidazole (Wang *et al.* 2015). Ainsi, un contrôle et une adaptation fine des conditions de culture sont essentiels pour l'obtention de grandes quantités de vitamine B12.

### 2.3. L'obtention d'acide propionique

L'acide propionique est un acide gras à chaîne courte, saturé, de masse moléculaire 74,08 g/mol, également connu sous le nom d'acide propanoïque ou C<sub>3</sub>H<sub>6</sub>O<sub>2</sub> (GESTIS Substance Database – disponible en [http://gestisen.itrust.de/nxt/gateway.dll/gestis\\_en/012590.xml?f=templates\\$fn=default.htm\\$3.0](http://gestisen.itrust.de/nxt/gateway.dll/gestis_en/012590.xml?f=templates$fn=default.htm$3.0)) (Figure 6).

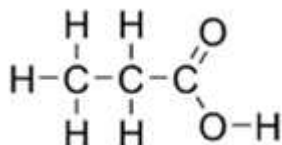


Figure 6 : Représentation de la formule chimique de l'acide propionique.

Cet acide est largement utilisé dans les industries suivantes: la pétrochimie, les produits cosmétiques, les plastiques, les herbicides et les agents de conservation. Dans ce contexte, la biologie synthétique est un secteur innovant, impliquant la biologie et l'ingénierie, qui permet de développer de nouvelles fonctions et de nouveaux systèmes, par exemple, l'utilisation d'une voie biochimique spécifique pour obtenir un matériau à travers une source d'énergie renouvelable (Parizzi *et al.* 2012).

La fermentation est l'un des procédés les plus utilisés dans l'industrie pour l'obtention de divers produits et sous-produits, parmi lesquels l'acide propionique. Actuellement, plusieurs stratégies visant l'obtention de ce produit ont été développées en utilisant différentes sources de carbone, telles que la fermentation du sucre de canne (saccharose), ces stratégies recherchent une réduction des coûts et une amélioration de la production de cet acide. Les bactéries propioniques sont capables de synthétiser l'acide propionique et elles représentent une source d'intérêt pour cet usage. Aussi, des études visant la physiologie et le métabolisme des Propionibacteries ont été réalisées afin, entre autres, d'identifier les meilleures espèces et les meilleures souches pour une utilisation industrielle (Thierry *et al.* 2011).

La fermentation de lactate, par le métabolisme initial de *P. freudenreichii*, génère des produits, tels que l'acétate, le propionate et le CO<sub>2</sub>, qui jouent des rôles variés et importants pour les industries alimentaires et de biotechnologie (Meurice *et al.* 2004). Par exemple, le propionate de calcium est un produit dérivé de l'acide propionique ayant la capacité d'agir comme agent de conservation, capable de lutter contre les bactéries défavorables et les champignons dans les aliments.

C'est à partir de ces observations sur l'acide propionique que Stowers et collaborateurs ont évalué 17 souches de *Propionibacterium* et ont défini des paramètres

optimaux de production et de rendement pour l'acide propionique en fermenteur. En suivant ces paramètres, des souches des espèces *P. acidipropionici* et *P. freudenreichii* produisent des quantités élevées d'acide propionique (Stowers, Cox and Rodriguez 2014).

Cependant, il est toujours difficile d'obtenir un bon rendement d'acide propionique à partir de Propionibacteria, car la génération de sous-produits dérivés de son métabolisme, tels que l'acide succinique, l'acide lactique, l'acide acétique réduisent le rendement de l'acide propionique (Stowers, Cox and Rodriguez 2014). Dans cette optique, il est important de mener des études pour augmenter l'efficacité de la production de ce produit. Cette production trouve aussi des débouchés en chimie verte à travers de la production de propylène ( $C_3H_6$ ), précurseur de l'acide propionique, un des hydrocarbures utilisés dans l'industrie pétrochimique. La polymérisation du propylène conduit au polypropylène, un composé largement utilisé dans diverses industries pour générer des carburants, des emballages jetables, des textiles, utilisés dans les appareils électro-ménagers et l'industrie automobile (Ammar *et al.* 2014). Ainsi, les bactéries qui synthétisent l'acide propionique à fort rendement font l'objet d'un intérêt croissant, en raison de l'essor de la chimie verte et d'un marché économique porteur pour les produits bios-sourcés.

## 2.4. Probiotique

Nous avons déjà évoqué que plusieurs bactéries, notamment des bactéries lactiques et propioniques sont l'objet d'intérêt en raison de leurs propriétés probiotiques. A ce titre, *P. freudenreichii* est largement étudié pour sa capacité à inhiber certains microorganismes indésirables, à moduler la microflore du côlon par la production de 1,4-dihydroxy-2-naphtoïque (DHNA), un facteur bifidogène, (Pérez Chaia, Zarate and Oliver 1999) et à survivre à différents stress (Isawa *et al.* 2002; Okada *et al.* 2013), notamment les stress rencontrés pendant le transit au sein du tube digestif (Jan *et al.* 2001; Leverrier *et al.* 2003a; Saraoui *et al.* 2013b). *P. freudenreichii* est également investigué à déclencher l'apoptose dans les cellules cancéreuses du côlon, à adhérer aux cellules épithéliales intestinales et au mucus, stimulant ainsi des activités immunomodulatrices et la prolifération cellulaire chez l'hôte (Ouwehand *et al.* 2002; Zárate *et al.* 2002). Dans cette section, nous présenterons une courte synthèse des études portant sur les diverses propriétés probiotiques attribuées à *P. freudenreichii*.

### 2.4.1 L'histoire et la définition du terme probiotique

Bien que les bénéfices de certains aliments fermentés pour la santé soient étudiés depuis longtemps (travaux de Metchnikoff au début du XX<sup>e</sup> siècle), l'origine du terme probiotique date de 1965, lorsque Lilly et Stillwell ont associé la survie de certains microorganismes à l'effet bénéfique de leurs produits (Lilly and Stillwell 1965). Actuellement, les probiotiques sont définis comme des microorganismes vivants qui, lorsqu'ils sont administrés d'une manière appropriée, apportent des effets positifs sur la santé de l'hôte (Bron, van Baarlen and Kleerebezem 2012). En raison de leur importance, les publications sur PubMed les concernant ont pris de l'ampleur au cours des dernières années (Figure 7).



## Les publications sur les probiotiques

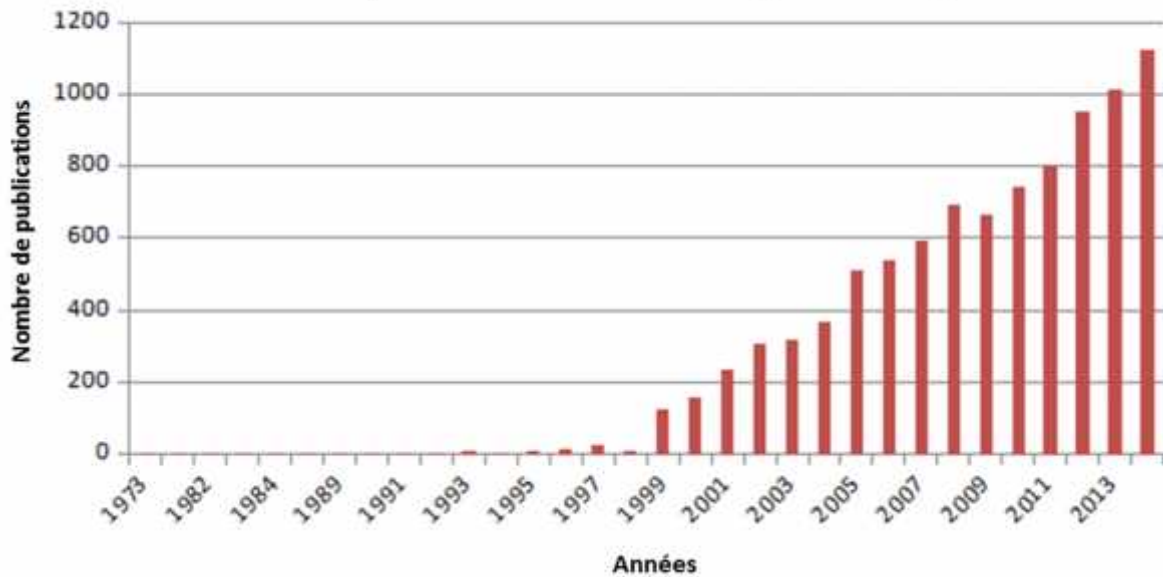


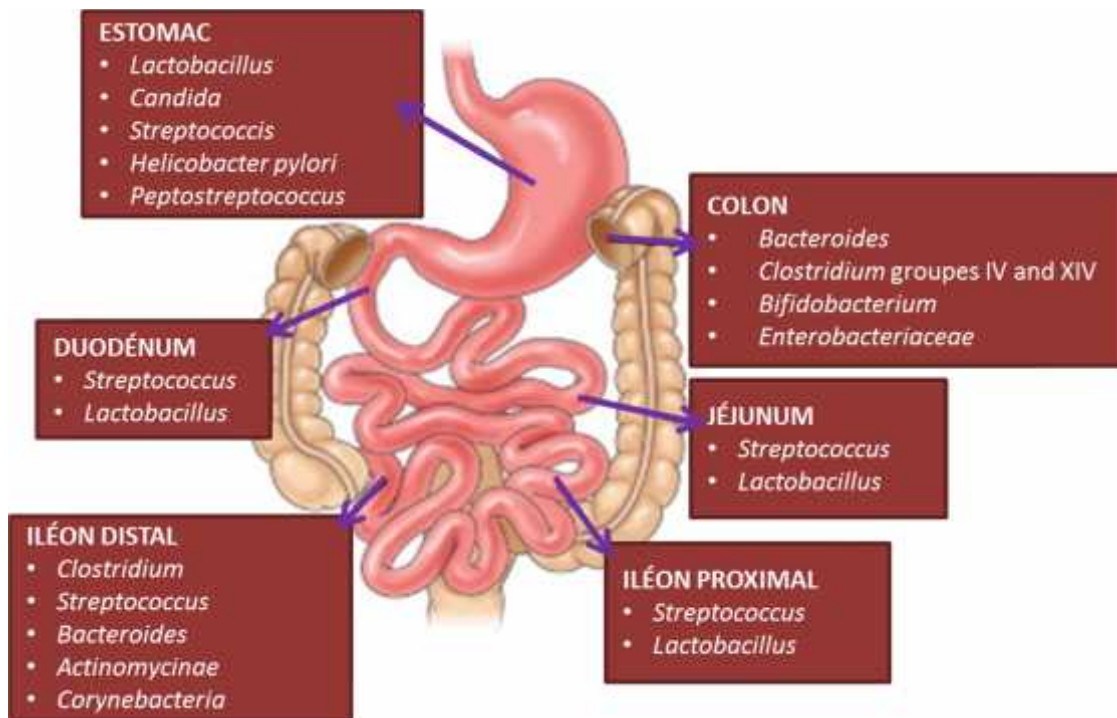
Figure 7 : Publications dans PubMed avec les mots clés « probiotics bacteria » au cours des dernières années.

Cependant, pour une utilisation des probiotiques chez l'homme et pour que des allégations santé soient acceptées, il est nécessaire d'effectuer des essais rigoureux *in vitro* et *in vivo*, qui visent à démontrer la qualité et la sécurité de ces organismes avant leurs applications dans la prévention ou le traitement de maladies et la supplémentation alimentaire. Quelques espèces bactériennes comme *Lactobacillus paracasei*, *L. johnsonii*, *L. fermentum*, *L. plantarum*, *L. rhamnosus*, *Bifidobacterium longum* et *P. freudenreichii* font l'objet d'étude sur leurs capacités à améliorer la santé humaine, notamment par le contrôle et la restauration de l'équilibre du microbiote intestinal (Parvez *et al.* 2006; Cousin *et al.* 2012; Santos Rocha *et al.* 2012; Luerce *et al.* 2014).

### 2.4.2 Contrôle du microbiote

Les bactéries propioniques peuvent avoir un effet sur la microflore des animaux et/ou des humains en inhibant des microorganismes indésirables et/ou en augmentant la prolifération des populations de *Bifidobacterium* sp., des bactéries qui jouent un rôle positif sur le tube digestif (Picard *et al.* 2005). Leur augmentation présente un certain nombre d'avantages physiologiques et cliniques pour la santé humaine et/ou animale, tels que : (i) accélérer le temps de passage à travers le tractus gastro-intestinal ; (ii) exercer un effet de barrière, car elles produisent des substances qui bloquent les sites de liaison de pathogènes ; (iii) exercer des effets sur le système immunitaire et dans les maladies du tractus gastro-intestinale ; (iv) réduire le risque de diarrhée lors d'infection à rotavirus ; (v) exercer des

effets sur la prolifération des cellules intestinales et (vi) empêcher la carcinogénèse du côlon (Picard *et al.* 2005). En plus des Bifidobactéries, il existe d'autres espèces bactériennes dominantes qui colonisent différentes parties du tractus gastro-intestinal et qui sont d'une grande importance pour le contrôle du microbiote intestinal. Ces espèces sont présentées dans la Figure 8.



**Figure 8 : Espèces microbiennes dominantes dans les différentes parties du tractus gastro-intestinal. Figure adaptée de l'article "Influences microbiennes dans les maladies inflammatoires chroniques de l'intestin" (Sartor 2008).**

Le microbiote intestinal joue un rôle important dans le système immunitaire chez les humains et/ ou les animaux et il est généralement acquis au cours de la première année qui suit la naissance. Il peut être influencé par des facteurs environnementaux externes, tels que la consommation de produits laitiers dans l'alimentation (Bernstein 2014; Geuking *et al.* 2014).

La composition du microbiote intestinal présente un intérêt fondamental de par ses liens avec certaines pathologies gastro-intestinales. De nombreuses études cherchent ainsi à améliorer notre compréhension de ces communautés microbiennes et de leur effet de modulation du système immunitaire et la santé de l'hôte en général (Sartor 2008; Hooper and Macpherson 2010; de Vos and de Vos 2012; Vangay *et al.* 2015; Velasquez-Manoff 2015).

### 2.4.3. Immunomodulation

L'immunomodulation est une propriété importante des bactéries probiotiques, permettant l'induction et la régulation de la réponse immunitaire. Elle est aussi utile pour la prévention ou le traitement du cancer, des maladies auto-immunes et dans le contrôle des microorganismes indésirables (Ouwehand et al. 1999; Sathish et al. 2013).

L'association de bactéries propioniques à des bactéries lactiques et des bifidobactéries peut avoir des effets bénéfiques sur des pathologies gastro-intestinales chez les enfants et les personnes âgées. Dans ce cadre, les bactéries propioniques agiraient indirectement en favorisant la croissance et le maintien à des concentrations élevées des bifidobactéries au cours de la période du traitement (Mantere-Alhonen 1995). Les bactéries propioniques seules semblent aussi pouvoir exercer un effet probiotique à travers une immunomodulation (Cousin *et al.* 2010; Le Maréchal *et al.* 2015). L'effet immunomodulateur des souches probiotiques est souvent associé à leur capacité d'adhésion à l'épithélium intestinal. Cette adhésion leur permet de se maintenir plus longtemps dans le tractus digestif, lors du transit. Cette capacité d'adhésion peut être un critère de sélection de probiotiques et certaines souches de *P. freudenreichii* ont fait l'objet d'études dans ce domaine (Ouwehand et al. 2002). L'implication des protéines de surface dans le rôle immunomodulateur de ces bactéries a été testée par la caractérisation protéomique du « surfaçome » de *P. freudenreichii*. Il a ainsi été montré que certaines protéines de surface, extraites de souches de *P. freudenreichii* et purifiées, étaient capables de reproduire, *in vitro*, une réponse anti-inflammatoire observée avec la souche complète (Le Maréchal et al. 2015).

#### 2.4.4. Apoptose induite par les acides gras à chaîne courte

Les bactéries propioniques sont connues pour leur capacité à produire des acides gras à chaîne courte (SCFA pour short chain fatty acids), tels que le propionate et l'acétate qui sont les principaux produits de fermentation de *P. freudenreichii*. Il a été observé que ces SCFA ont des effets intéressants sur la physiologie des cellules épithéliales intestinales (Jan *et al.* 2002; Pérez Chaia and Zarate 2005).

Les SCFA, et principalement le butyrate, constituent une source d'énergie privilégiée, un véritable carburant, pour les cellules épithéliales intestinales. En 1995, Hague et collaborateurs ont montré que ces mêmes SCFA, par exemple, le butyrate de sodium, sont également capables au contraire d'induire l'apoptose dans les cellules cancéreuses du colon (Hague *et al.* 1995). Par la suite, les travaux de Jan et collaborateurs en 2002, ont montré *in vitro* que des souches de *P. acidipropionici* et *P. freudenreichii* sont elles aussi capables d'induire l'apoptose dans des lignées de cellules humaines de cancer colorectal, telles que HT29 et Caco2. Cet effet a aussi été observé sur des lignées cellulaires humaines de cancer du col de l'utérus (HeLa). Les SCFA induisent des signes caractéristiques de l'apoptose, tels que : (i) la perte du potentiel transmembranaire des mitochondries, (ii) la génération d'espèces réactives de l'oxygène, (iii) d'induction de la cascade caspase-3 et (iv) la condensation de la chromatine au niveau nucléaire. Le mécanisme général de l'apoptose via les SCFA est schématisé sur la Figure 9. Par ailleurs, cette activité pourrait s'exercer *in vivo* car ces souches restent vivantes et métaboliquement actives dans l'intestin humain (Jan *et al.* 2002).

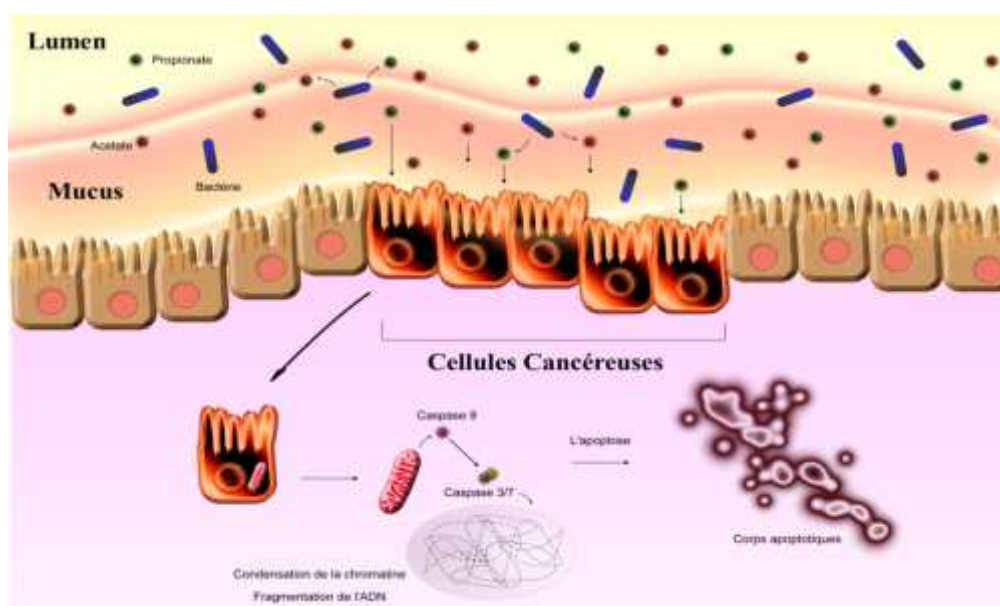


Figure 9 : Induction de l'apoptose à partir de SCFAs produites par les bactéries propioniques dans le côlon.

Il a ainsi été proposé que les bactéries propioniques pourraient venir compléter l'arsenal d'options thérapeutiques disponibles dans la prévention ou le traitement du cancer du colon (Lan *et al.* 2008). Plus récemment, la même équipe a proposé une nouvelle formulation d'un lait fermenté avec certaines souches de *P. freudenreichii*, visant une alimentation préventive ou un supplément dans l'alimentation de personnes qui sont prédisposées à ce type de cancer (Cousin *et al.* 2012).

#### **2.4.5. Molécule acide 1,4-dihydroxy-2-naphtoïque (DHNA)**

Le composant 1,4-dihydroxy-2-naphtoïque (DHNA) présente un effet bifidogène marqué (i.e., il favorise la croissance des Bifidobactéries) et constitue de ce fait un prébiotique. Les prébiotiques sont généralement des oligosaccharides non digérés par les enzymes digestives chez l'homme. Ainsi, la croissance de *P. freudenreichii*, qui produit du DHNA, influe directement sur la croissance des bifidobactéries, faisant de *P. freudenreichii* une bactérie combinant des activités pro- et pré-biotiques. Elle contribue ainsi à l'équilibre de la flore intestinale, en favorisant le développement de bactéries bénéfiques. Le DHNA permettrait par ailleurs de réduire l'inflammation du colon (Quigley, Hudson and Englyst 1999; Okada *et al.* 2006). *In vivo*, l'administration de DHNA entraîne aussi des modifications du système immunitaire intestinal chez la souris, en modifiant la réponse immunitaire innée chez les cellules intestinales (Okada *et al.* 2013).

De même, le DHNA est une molécule prometteuse qui pourrait être utilisée dans le traitement de la colite. Dans un modèle de colite induite chez la souris, son administration entraîne une amélioration dans l'histologie de l'épithélium intestinal, une suppression de la migration des lymphocytes et de la synthèse d'interleukines pro-inflammatoires (Okada *et al.* 2013). Dans ce contexte, des recherches supplémentaires doivent être effectuées visant à une optimisation du DHNA pour des applications thérapeutiques et pour montrer que les quantités produites *in situ* par *P. freudenreichii* pourraient aboutir aux mêmes effets.

#### 2.4.6. Capacité d'adaptation

Cette sous-section fait le point sur la capacité de survie de *P. freudenreichii* et sa capacité d'adaptation à différents stress. Elle sera divisée en deux parties: la tolérance au stress thermique et au stress digestif. La tolérance à ces stress est une propriété très importante chez les bactéries propioniques pour leur maintien et leur viabilité, ce qui optimise leur utilisation comme probiotiques (Leverrier *et al.* 2003a).

##### Ñ Tolérance au stress thermique:

Les bactéries propioniques survivent à une large gamme de températures. Dans les étapes d'affinage des fromages PPC, elles résistent autant à basse température (4°C), qu'à moyenne (30°C) et haute température (52-53°C). Ces températures varient en fonction des différentes étapes du processus de production et du type de fromage. Pour le fromage Gruyère traditionnel, il y a une étape de cuisson soumettant les bactéries à une température très élevée, entre 52-53°C, pendant 30 minutes. La thermotolérance à 55°C de différents souches utilisées dans la fabrication de Gruyère a été étudiée par une approche protéomique et les profils obtenus en électrophorèse 2D ont révélé différentes réponses selon que les souches sont thermotolérantes ou thermosensibles (Anastasiou *et al.* 2006). Dans les espèces thermotolérantes, les protéines associées au processus de la fermentation propionique, au métabolisme des acides aminés, aux nucléotides et à la phosphorylation oxydative étaient surexprimés. En particulier, les chaperons moléculaires et les composés antioxydants de type thiol sont induits par la chaleur. Toutefois, il a aussi été observé que, lors de l'adaptation thermique, les voies de la transduction du signal, la biosynthèse et le maintien de la paroi cellulaire contribuent à maintenir le processus de fermentation (Anastasiou *et al.* 2006). D'autres études portant sur les bactéries propioniques ont également rapporté que les chaperons moléculaires et des protéines de choc thermique sont induites dans de nombreuses situations de stress (Jan *et al.* 2001; Leverrier *et al.* 2003a). Cette différence dans l'expression des protéines des diverses souches de bactéries propioniques renforce l'argument selon lequel il y a différents phénotypes au sein de la même espèce (Anastasiou *et al.* 2006). Cette forme de tolérance est importante chez *P. freudenreichii* pour maintenir la survie bactérienne et assurer le bon déroulement du processus de fabrication des fromages de type PPC. Elle constitue en outre un avantage indéniable pour son utilisation comme probiotique, si on y associe sa résistance aux stress digestifs (Leverrier *et al.* 2003b).

## Ñ Tolérance au stress digestif:

Pour qu'un probiotique exerce sa fonction, il faut maintenir son activité métabolique et sa survie après son passage à travers le tractus gastro-intestinal. Le travail de Lan et collaborateurs a démontré cette capacité chez *P. freudenreichii* en démontrant la viabilité des bactéries dans le tractus digestif et dans le colon de la souris (Lan *et al.* 2007). Cette étude a également confirmé la tolérance au pH acide de certaines souches spécifiques de *P. freudenreichii*, comme cela avait déjà été rapporté dans les études précédentes portant sur la résistance au stress acide (Jan, Rouault and Maubois 2000), mais aussi sur la tolérance aux sels biliaires (Warminska-Radyko, Laniewska-Moroz and Babuchowski 2002). Des expériences de protéomique, utilisant l'électrophorèse 2D, ont montré 24 protéines induites pendant l'adaptation de *P. freudenreichii* dans un milieu contenant des sels biliaires. On compte parmi elles des protéines de transduction du signal et des facteurs sigma alternatifs.

Pour évaluer l'activité métabolique de *P. freudenreichii*, une RT-PCR quantitative utilisant une sonde *TaqMan* et ciblant l'ARNm codant pour l'enzyme transcarboxylase a été développée. Cette méthode non-invasive et efficace a été utilisée pour prouver la survie et l'activité métabolique de *P. freudenreichii* dans le tractus intestinal des souris (Hervé *et al.* 2007). D'autres méthodes cherchant à évaluer l'activité métabolique de *P. freudenreichii* dans le fromage ont aussi été développées dans le but de caractériser ces souches pour leur survie. Dans ce contexte, une méthode de RT-PCR quantitative utilisant du *SyberGreen* a été optimisée pour la quantification de cellules et d'ARNm marqueurs (gènes 16S, tuf, groL) de *P. freudenreichii* et de *Lactobacillus paracasei*, pendant le processus de production des fromages PPC (Falentin *et al.* 2010). Cette méthode a permis de quantifier la viabilité cellulaire, et d'aller plus loin dans les études concernant la survie de *P. freudenreichii*.



### 3. Génomique de *Propionibacterium freudenreichii*

Grâce aux technologies de séquençage de nouvelle génération (NGS), il a été possible d'obtenir le séquençage de plusieurs génomes de bactéries propioniques, ce qui a conduit à une meilleure connaissance biologique de l'espèce. Ces génomes présentent, en général, une teneur élevée en GC, environ 67% (Falentin *et al.* 2010a) et certaines souches peuvent héberger en outre des plasmides (Luijk *et al.* 2002). Les données provenant du séquençage peuvent être utilisées dans les études comparatives et/ou évolutives visant à comprendre la génomique, au niveau structurel et au niveau fonctionnel de ce microorganisme.

En 2010, le STLO a publié le premier génome complet de *P. freudenreichii* (souche CIRM-BIA1) qui a été séquencée en utilisant la méthode traditionnelle de Sanger (Falentin *et al.* 2010a). L'assemblage initial de ce génome a été effectué sans séquences répétitives, en utilisant le logiciel Phred/Phrap/CONSED ([www.phrap.org](http://www.phrap.org)) (Vallenet *et al.* 2008; Falentin *et al.* 2010a). Ensuite, pour la clôture des *gaps*, il a été nécessaire d'utiliser le *kit sequence GC finishing* (GEhealthcare) et le *kit transposition bombs* (Falentin *et al.* 2010a). Ce génome présente un chromosome circulaire de 2,7 Mb, de haute teneur en GC (67%) et est disponible dans la base de données *Genbank*, avec le numéro d'accès FN806773 (Falentin *et al.* 2010a).

Depuis, 21 autres génomes de *P. freudenreichii* ont été séquencés par l'équipe. Leur séquençage a été réalisé par la méthode Illumina, en utilisant la bibliothèque *paired-ends*, sauf pour la souche CIRM-BIA 9 qui a utilisé la bibliothèque *single reads*. L'assemblage de ces 21 génomes a été mené par une approche *de novo*, puis les *scaffolds* ont été ordonnés par comparaison avec le génome de référence complet CIRM-BIA1 (Falentin *et al.* 2010a; Loux *et al.* 2015a). Les détails de l'assemblage de ces 21 génomes sont présentés dans le Tableau 5. Tous les génomes ont été annotés sur la plate-forme AGMIAL, qui est une interface graphique pour l'annotation de génomes complets et incomplets des procaryotes. Les 21 génomes ont été déposés dans la base de données de EBI-EMBL et ces données sont disponibles en suivant le lien : <http://www.ebi.ac.uk/ena/data/view/Taxon:Propionibacterium%20freudenreichii> (Loux *et al.* 2015b). Ces 21 génomes comprennent notamment celui *P. freudenreichii* CIRM-BIA138, une des souches utilisées au cours de ces travaux de thèse pour l'analyse en RNA-Seq.

Cependant, la plupart de ces nouveaux génomes de *P. freudenreichii* sont encore incomplets, à l'état de *drafts*, cela est dû à la difficulté de l'étape d'assemblage liée à la présence de nombreuses séquences répétées (transposons).

Des analyses bioinformatiques ont été réalisées sur ces différents génomes de *P. freudenreichii*, afin d'identifier les gènes impliqués dans (i) la résistance aux différents pH et à la présence des sels biliaires; (ii) la synthèse des vitamines et des acides aminés et (iii) le stockage du carbone. Ces analyses montrent qu'il s'agit d'une bactérie versatile, dotée d'un large arsenal de gènes pour survivre dans des environnements défavorables (Falentin *et al.* 2010a; Loux *et al.* 2015b). Ces données contribuent à expliquer la capacité de *P. freudenreichii* à maintenir une activité métabolique en conditions de stress et de survie à long terme. Ceci explique pour partie sa viabilité dans des environnements qui simulent le tractus gastro-intestinal chez les humains et les animaux. Ces analyses génomiques ont aussi été largement utilisées pour identifier des gènes potentiellement liés à son activité immunomodulatrice.

D'autres études ont révélé la présence de mutations ponctuelles et d'îlots génomiques, variables d'une souche à l'autre. Ceci montre la plasticité chromosomique de cette espèce, et résulte probablement de transferts horizontaux, duplications, transpositions et mutations accumulées. L'analyse de ces variations inter-souches, couplée à une caractérisation phénotypique complète de leur capacité à métaboliser les sucres a permis d'établir un lien entre cette plasticité et le fait que certaines souches de *P. freudenreichii* utilisent, préférentiellement, certaines sources de carbone. Ces analyses ouvrent aussi de nouvelles perspectives pour améliorer leur sélection et leur utilisation pour l'industrie fromagère et des probiotiques (Loux *et al.* 2015b).

Le génome complet de *P. freudenreichii* CIRM-BIA 1 a aussi été analysé et comparé à celui de *P. acnes* afin d'identifier les facteurs différenciant ces deux espèces proches. Cette analyse a aussi été menée pour cerner ce qui sous-tend les capacités probiotiques de certaines souches de *P. freudenreichii* (Falentin *et al.* 2010a).

Ces analyses génomiques structurales basées sur l'ADN sont complétées d'analyses fonctionnelles à travers des études portant sur l'ARN, les protéines et les produits de métabolisme de cette bactérie (respectivement, des approches transcriptomique, protéomique et métabolomique). Cela est illustré dans le deuxième chapitre de la section de résultats de ce manuscrit thèse : l'utilisation de la génomique fonctionnelle pour comprendre le phénomène de la survie à long terme chez *P. freudenreichii*.

**Tableau 5 : Génomes de *P. freudenreichii* déposés par le STLO.**

Number in CIRM-BIA collection (Rennes, France)	Locus_tag	Isolated from (date)	Bioproject	Raw reads accession number	Read quantity	Read length	Sequencing depth	Number of scaffolds >1000 bp	Cumulative contig length	N50	Number of N segments	Number of complete genes
9	PFCIRM9	Emmental cheese (1989)	<b>PRJEB6427</b>	ERS638416	26 013 151	37	370	204	2 533 825	21 887	0	2 308
118	PFCIRM118	Gruyère cheese (1973)	<b>PRJEB6428</b>	ERS638410	51 926 392	37	739	41	2 583 496	193 767	59	2 325
119	PFCIRM119	Gruyère cheese (1973)	<b>PRJEB6430</b>	ERS638409	58 084 132	37	827	46	2 644 402	128 788	91	2 422
121	PFCIRM121	Swiss cheese (1937)	<b>PRJEB6431</b>	ERS638414	52 432 380	37	746	59	2 561 530	86 244	68	2 304
122	PFCIRM122	NCIB (1992)	<b>PRJEB6432</b>	ERS638421	54 950 090	37	782	78	2 614 705	73 741	70	2 348
123	PFCIRM123	Morbier cheese (1992)	<b>PRJEB6438</b>	ERS638423	54 523 668	37	776	47	2 604 827	114 519	78	2 377
125	PFCIRM125	Emmental cheese (1992)	<b>PRJEB6434</b>	ERS638418	54 868 488	37	781	43	2 511 448	238 430	59	2 272
127	PFCIRM127	Emmental cheese (1992)	<b>PRJEB6435</b>	ERS638428	55 178 920	37	785	57	2 567 344	81 280	53	2 330
129	PFCIRM129	Emmental cheese (1992)	<b>PRJEB4826</b>	ERS638425	52 269 870	37	744	59	2 591 314	123 180	59	2 338
134	PFCIRM134	NA	<b>PRJEB6441</b>	ERS638419	50 523 582	37	719	47	2 635 396	153 798	152	2 318
135	PFCIRM135	ewe raw milk (1994)	<b>PRJEB6442</b>	ERS638422	50 649 408	37	721	47	2 620 043	114 923	75	2 343
138	PFCIRM138	Emmental cheese (1992)	<b>PRJEB6433</b>	ERS638424	58 100 372	37	827	72	2 573 580	80 515	110	2 304
139	PFCIRM139	Emmental cheese (1992)	<b>PRJEB6436</b>	ERS638426	60 700 312	37	864	45	2 589 117	106 963	53	2 362
456	PFCIRM456	raw milk Raclette cheese (1992)	<b>PRJEB6445</b>	ERS638427	49 094 398	37	699	38	2 502 068	131 281	61	2 253
508	PFCIRM508	Gruyère (1973)	<b>PRJEB6429</b>	ERS638411	50 259 834	37	715	47	2 578 004	141 315	93	2 302
512	PFCIRM512	raw milk Morbier cheese (1994)	<b>PRJEB6439</b>	ERS638413	63 808 750	37	908	45	2 580 419	153 592	57	2 349
513	PFCIRM513	egyptian ras cheese (1995)	<b>PRJEB6440</b>	ERS638415	62 230 134	37	886	48	2 659 761	179 407	54	2 416
514	PFCIRM514	hay (1994)	<b>PRJEB6443</b>	ERS638417	51 481 488	37	733	51	2 521 626	121 592	66	2 285
516	PFCIRM516	nepalian yack cheese (2000)	<b>PRJEB6444</b>	ERS638420	61 823 266	37	880	62	2 630 828	125 877	205	2 342
527	PFCIRM527	Fribourg cheese (1992)	<b>PRJEB6437</b>	ERS638412	69 806 362	37	993	63	2 576 369	104 864	57	2 344
1025	PFCIRM1025	Emmental cheese (1992)	<b>PRJEB6446</b>	ERR738371	61 323 626	81	1 910	104	2 699 685	62 523	19	2 308

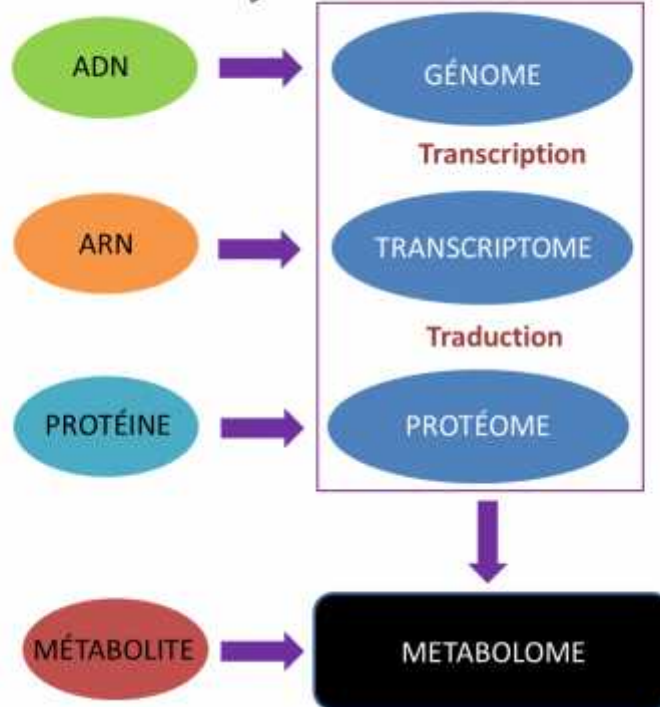
#### 4. Transcriptomique de *Propionibacterium freudenreichii*

Au cours des 10 dernières années, plusieurs études ont porté sur l'expression des gènes chez *Propionibacterium* afin de mieux comprendre et d'expliquer comment se produit leur régulation dans ce groupe bactérien. Ces données transcriptomiques sont déposées dans des bases, comme l'EMBL-EBI (<http://www.ebi.ac.uk/genomes/>), créées pour permettre l'accès public, qui fournissent des données du génome et du transcriptome de différentes bactéries. Ces travaux nous aident à comprendre comment les bactéries propioniques s'adaptent et régulent l'expression des gènes en réponse à différentes conditions environnementales.

Des espèces comme *P. acnes* et *P. freudenreichii* sont d'une grande importance dans ce genre, car *P. acnes* est commensale et fait partie de la flore de la peau ; quant à *P. freudenreichii*, nous avons déjà mentionné plus haut qu'elle est d'intérêt pour l'industrie laitière dans le secteur des fromages PPC et d'intérêt biotechnologique pour la synthèse de biomolécules et pour ses propriétés probiotiques. Les premières études de transcriptomiques sur *P. freudenreichii* ont été réalisées avec les techniques de puce d'ADN et RT-qPCR. Ces études ont montré la survie et l'adaptation de ces bactéries à des températures variées pendant le processus d'affinage des fromages PPC, ainsi que leur adaptation métabolique (*in situ*) dans le côlon de porcs (Dalmaso *et al.* 2012b; Saraoui *et al.* 2013b). Afin d'accroître les connaissances sur la survie de cette bactérie, nous avons utilisé le RNA-Seq, une approche transcriptomique plus exhaustive, couplé à une analyse métabolomique ciblée. Nous présenterons, dans la section II des résultats de ce manuscrit, un article concernant une analyse de la souche *P. freudenreichii* CIRM-BIA138, intégrant ses données transcriptomiques et métabolomiques obtenues *in vitro*.

L'association de ces données aux niveaux ARN et biochimique, aux données protéomiques déjà obtenues concernant *P. freudenreichii* permettra d'aborder l'étude fonctionnelle de cette bactérie et nous éclairera sur ses capacités d'adaptation et de survie (Figure 10).

**Corrélation des données génomiques, transcriptomiques et protéomiques de *P. freudenreichii***



**Figure 10 : Combinaison des données des données génomiques, transcriptomiques et protéomiques.**

## 5. Les principales voies métaboliques de *Propionibacterium freudenreichii*

Les bactéries propioniques tirent leur énergie de divers substrats carbonés par la fermentation propionique. Chez *P. freudenreichii*, l'association des cycles de Krebs et de Wood-Werkman sont caractéristiques de l'espèce.

Le cycle de Krebs, appelé également cycle de l'acide citrique ou cycle de l'acide tricarboxylique (TCA, en anglais), correspond à des réactions métaboliques qui se produisent dans le cytoplasme des cellules bactériennes. Ce cycle est une voie amphibolique présentant des étapes de (i) dégradation des métabolites pour générer de l'énergie – ATP – (voie catabolique) et (ii) production de précurseurs pour la biosynthèse (voie anabolique) (Nelson *et al.* 2008).

Chez *P. freudenreichii*, le cycle Wood-Werkman, également connu sous le nom de voie de carbone central, est une composante centrale du métabolisme. Ce cycle permet la formation de propionate à partir du pyruvate (Wood 1981). Au cours de la fermentation, les bactéries propioniques sont capables d'utiliser le lactate et de produire, comme cité précédemment, certains sous-produits d'intérêt pour la composition de l'Emmental. La Figure 11 illustre le métabolisme de *P. freudenreichii*, basé sur le cycle Wood-Werkman et son association au cycle de Krebs.

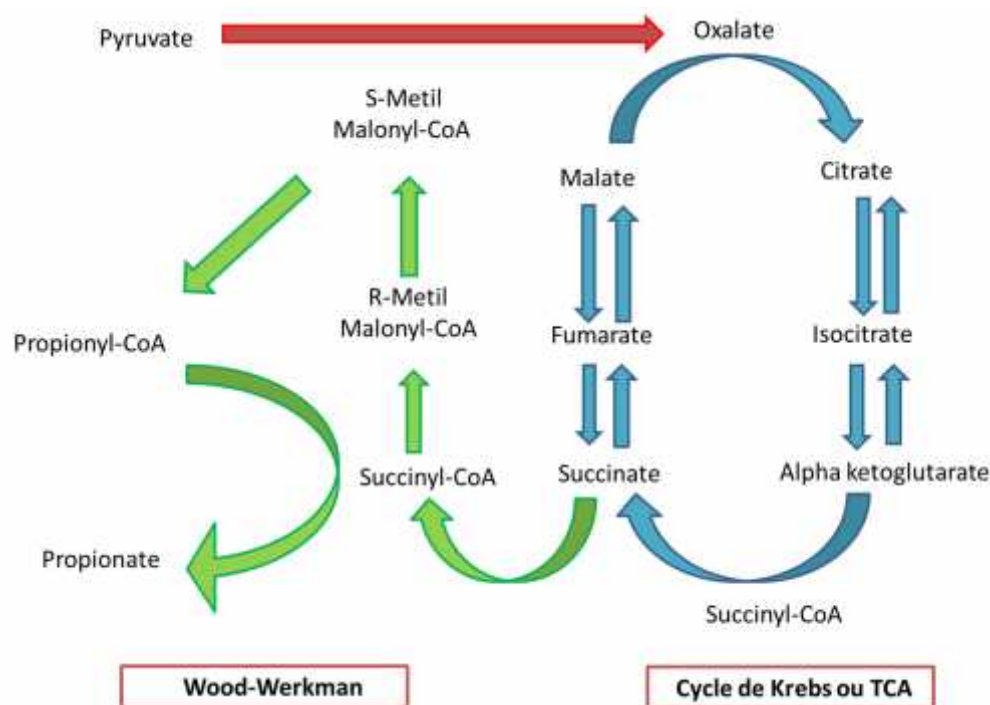
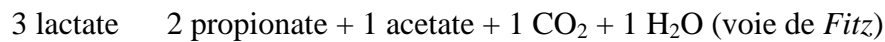
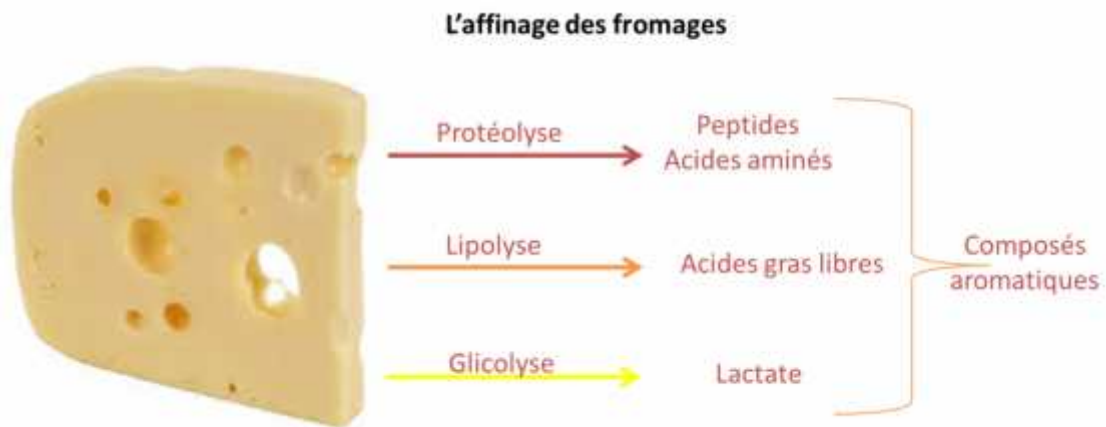


Figure 11 : Voie métabolique de Wood-Werkman chez *P. freudenreichii* liée au cycle de Krebs. Figure adaptée de l'article: A Temporal -omic Study of *Propionibacterium freudenreichii* CIRM-BIA1<sup>T</sup> Adaptation Strategies in Conditions Mimicking Cheese Ripening in the Cold (<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0029083>).

La réaction globale décrite dans ces bactéries suit l'équilibre suivant: (Crow 1986)



*P. freudenreichii* jouant un rôle majeur dans la maturation des fromages PPC et le développement de leur flaveur (Figure 12). Aussi les diverses voies métaboliques aboutissant à la formation de composés d'arôme ont-elles été recherchées, notamment en combinant des études génomiques et métabolomiques. Diverses voies métaboliques (métabolisme de la glycolyse, de la lipolyse, des acides aminés, de stockage d'énergie) ont ainsi été reconstruites intégralement, pour certaines souches. Ces travaux ont fait l'objet de nombreuses articles et de revues détaillées (Thierry *et al.* 2011).



**Figure 12 : Schéma des processus métaboliques impliqués dans l'affinage des fromages PPC visant l'obtention des composés aromatiques.**

## **CAPÍTULO 2 : RESULTADOS**

## **CHAPITRE 2 : RÉSULTATS**



**SEÇÃO I: Caracterização do fenótipo de  
linhagens de *Propionibacterium  
freudenreichii***

**SECTION I : Caractérisation du  
phénotype des souches de  
*Propionibacterium freudenreichii***

## 1. Introdução

A primeira seção de resultados será apresentada na forma de artigo científico. O objetivo deste estudo foi avaliar a variabilidade entre as linhagens e de caracterizar os fenótipos das células de *P. freudenreichii* durante a sobrevivência a longo prazo.

Este manuscrito descreve, inicialmente, a caracterização fenotípica de 8 linhagens de *P. freudenreichii*, em meio de cultura sem a adição de nutriente durante um período de 11 dias, sendo os 8 últimos correspondentes à uma carência nutricional. Durante este período, cada linhagem foi monitorada por densidade óptica, enumeração de colônias (ágar) e observação morfológica das colônias. A viabilidade e integridade da membrana também foram avaliadas. Para isso, foi utilizado o *kit live/dead* acompanhado de uma observação em microscopia de epifluorescência, e nós verificamos a ausência de lise celular qPCR com primers específicos para o gene alvo *tuf*.

Após 2 dias de cultura a população atingiu  $9 \log_{10}$  UFC/mL. Após 11 dias as 8 linhagens de *P. freudenreichii*, mantiveram uma população elevada, de aproximadamente  $7.8$  a  $8.2 \log_{10}$  UFC/mL. O lactato, principal composto de carbono foi totalmente consumido entre 2 e 3 dias após a inoculação, e as bactérias tiveram que lidar com um período de estresse nutricional. A integração dos métodos utilizados permitiu comparar os níveis de sobrevivência e serviu como base para a triagem das 8 linhagens. A linhagem CIRM-BIA 138 é aquela que tem mostrado uma melhor sobrevivência à 11 dias após a inoculação. As pesquisas desta tese foram então focadas nesta linhagem.

Existem três principais mecanismos que permitem a sobrevivência bacteriana, quando as bactérias estão expostas à condições ambientais desfavoráveis: (i) a esporulação, (ii) a morte celular programada e (iii) dormência. Já está estabelecido que *P. freudenreichii* não esporula. Quanto ao mecanismo de morte celular programada, os restos celulares de parte da população lisada fornecem nutrientes para a subpopulação não lisada, permitindo assim que determinadas espécies bacterianas enfrentem as condições de carência nutricional. Os resultados de qPCR mostram uma estabilidade da população bacteriana à  $9 \log_{10}$  e nos levam a concluir que *P. freudenreichii* CIRM-138 não usa a morte celular como mecanismo para a sobrevivência a longo prazo para manter a sua viabilidade, uma vez que a lise celular não é detectável nestas culturas.

Portanto, nossa hipótese é que esta bactéria permanece viável por causa da sua entrada em estado de dormência. Os resultados da microscopia de epifluorescência mostraram que metade das células bacterianas de todas as linhagens apresentaram

membranas permeabilizadas após 11 dias de cultura. Durante o período de sobrevivência, uma perda de  $1 \log_{10}$  UFC/mL foi observada em comparação aos resultados de quantificação da população por qPCR. A comparação dos resultados de contagem de colônias e de qPCR, assim como a ausência de lise sugerem que a 11 dias de cultura, 90% da população bacteriana perderam a capacidade de formar colônias em ágar, mas que uma boa parte das células são, no entanto, viáveis. A integração dos resultados da caracterização fenotípica sugerem que esta bactéria atinge um estado VBNC no momento o qual começa o período de sobrevivência a longo prazo. Este estado é descrito em muitas bactérias Gram-positivas e Gram-negativas, sendo uma estratégia para a sobrevivência nos períodos de estresse, tais como a temperatura, o pH e a disponibilidade de nutrientes variados.

Este trabalho mostra pela primeira vez um mecanismo de dormência abordado em *P. freudenreichii*, cujas células entram em estado VBNC para garantirem sua sobrevivência a longo prazo em condições de carência nutricional. Isso aumenta o nosso conhecimento nesta espécie de lacteos e abre perspectivas para novos estudos. Novas abordagens integradas, baseadas em técnicas *omics* devem ser realizadas para compreender o mecanismo de dormência e a regulação metabólica bacteriana durante este período de sobrevivência a longo prazo em *P. freudenreichii* (ver secção - Secção II deste manuscrito de tese).

## 1. Introduction

La première section des résultats sera présentée sous forme d'article scientifique. L'objectif de cette étude a été d'évaluer la variabilité inter-souches et de caractériser les phénotypes des cellules de *P. freudenreichii* pendant la survie à long terme.

Ce manuscrit décrit d'abord la caractérisation phénotypique de 8 souches de *P. freudenreichii*, en milieu de culture sans ajout sur une période de 11 jours dont les 8 derniers correspondent à une carence nutritionnelle. Pendant cette période, chaque souche a fait l'objet d'un suivi par densité optique, par dénombrement sur gélose et en observant la morphologie des colonies. La viabilité et l'intégrité membranaire ont aussi été évaluées. Pour cela, nous avons utilisé le *kit live/dead* suivi d'une observation en microscopie à épifluorescence et nous avons vérifié l'absence de lyse cellulaire par PCR quantitative en ciblant le gène *tuf*.

Après 2 jours de culture la population atteignait  $9 \log_{10}$  UFC/mL. Après 11 jours les 8 souches de *P. freudenreichii*, maintenaient une population élevée, d'environ  $7.8$  à  $8.2 \log_{10}$  UFC/mL. Le lactate, principal composé carboné étant épuisé entre 2 et 3 jours après inoculation, et les bactéries ont dû faire face à une période de stress nutritionnel. L'intégration des méthodes utilisées a permis de comparer les niveaux de survie et a servi de base au criblage des 8 souches. La souche CIRM-BIA 138 est celle qui a montré la meilleure survie à 11 jours après inoculation. Les recherches de la thèse se sont ensuite focalisées sur cette souche.

Il y a trois mécanismes majeurs qui permettent la survie bactérienne lorsque les bactéries sont exposés à des conditions environnementales défavorables : (i) la sporulation, (ii) la mort cellulaire programmée et (iii) la dormance. Il est déjà établi que *P. freudenreichii* ne sporule pas. Lors du mécanisme de mort programmée, les débris cellulaires d'une partie de la population lysée fournissent des nutriments pour la sous-population non-lysée, et permettent à certaines espèces bactériennes de faire face à des conditions de carence nutritionnelle. Les résultats de PCRq montrant une stabilité de la population bactérienne à  $9 \log_{10}$  nous amènent à conclure que *P. freudenreichii* CIRM-138 n'utilise pas la mort cellulaire programmée comme mécanisme de survie à long terme pour maintenir sa viabilité, car aucune lyse cellulaire n'est détectable dans les cultures.

Nous émettons donc l'hypothèse que cette bactérie reste viable en raison de son entrée en état de dormance. Les résultats de la microscopie à épifluorescence ont démontré que la moitié des cellules bactériennes de toutes les souches présentait des membranes perméabilisées après 11 jours de culture. Pendant la période de survie, une perte de  $1 \log_{10}$  UFC/mL a été observée par rapport aux résultats de quantification de la population par qPCR. La comparaison des résultats de dénombrement et de qPCR, ainsi que l'absence de lyse suggèrent qu'à 11 jours de culture, 90 % de la population bactérienne a perdu sa capacité à former des colonies sur gélose mais qu'une bonne partie des cellules sont néanmoins viables. L'intégration des résultats de la caractérisation phénotypique suggère que cette bactérie atteint un état VBNC au moment où commence la période de survie à long terme. Cet état est décrit chez de nombreuses bactéries à Gram positif et à Gram négatif, comme étant une stratégie pour la survie dans des périodes de stress tels que la température, le pH et la disponibilité de nutriments variés.

Ce travail montre pour la première fois qu'un mécanisme de dormance se met en place chez *P. freudenreichii*, dont les cellules entrent en état en VBNC pour assurer leur survie à long terme en condition de carence nutritionnelle. Cela accroît nos connaissances sur cette espèce laitière et ouvre des perspectives pour de nouvelles études. De nouvelles approches intégrées, basées sur des techniques *omiques* devraient être menées pour comprendre le mécanisme de dormance et la régulation métabolique bactérienne au cours de cette période de survie à long terme chez *P. freudenreichii* (voir article - section II de ce manuscrit de thèse).

**2. Artigo 1: The long-term survival of *Propionibacterium freudenreichii* in a context of nutrient shortage.**

Short running head: *P. freudenreichii* survival

**Flavia Figueira Aburjaile<sup>1,2,3</sup>, Marie-Noëlle Madec<sup>2,3</sup>, Sandrine Parayre<sup>2,3</sup>,  
Anderson Miyoshi<sup>1</sup>, Vasco Azevedo<sup>1</sup>, Yves Le Loir<sup>2,3†</sup>, Hélène Falentin<sup>2,3†</sup>**

<sup>1</sup> Department of General Biology, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

<sup>2</sup> INRA, UMR 1253, Science et Technologie du Lait et de l'Oeuf, 35000, Rennes, France

<sup>3</sup> AGROCAMPUS OUEST, UMR1253, UMR Science et Technologie du Lait et de l'Oeuf, 35000, Rennes, France

Correspondance: Hélène Falentin, INRA, UMR 1253, Science et Technologie du Lait et de l'OEuf, Rennes, France, e-mail: [helene.falentin@rennes.inra.fr](mailto:helene.falentin@rennes.inra.fr).

†: YLL and HF share credit as senior co-authors

## **Abstract**

**Aims:** *Propionibacterium freudenreichii* is an actinobacterium widely used in dairy industry during the ripening process of Swiss-type cheeses and which presents probiotic properties. PF is reportedly a hardy bacterium, able to survive during the cheese-making process and when subjected to digestive stresses. During this study the long-term survival of *P. freudenreichii* was investigated for 11 days by means of phenotypic characterization in a culture medium without the addition of any nutrients.

**Methods and Results:** For 11 days, in a non-nutrient supplemented culture medium, eight strains were monitored by measuring their optical density, counting colony-forming units (CFU) and using LIVE/DEAD staining and microscopy observation. Under these conditions, all strains displayed high survival rates in the culture medium, their culturability reaching more the  $9 \log_{10}$  CFU/mL after two days. After 11 days, this value ranged from 7.8 to  $8.2 \log_{10}$  CFU/mL depending on the strain, and at least 50% of the *P. freudenreichii* population displayed an intact envelope. Since lysis of part of a bacterial population may be a microbial strategy to recover nutrients, in CIRM-BIA 138 (the strain with the highest population at day 11), cell lysis was assessed by quantifying intact bacterial cells using qPCR targeting the housekeeping gene *tuf*. No lysis was observed.

**Conclusion:** Taken together, our results suggest that *P. freudenreichii* strains use a viable but nonculturable (VBNC) state to adapt to the long term survival phase.

**Significance and Impact of the study:** Assessing the viability of *P. freudenreichii* and understanding their mechanisms for survival should be of great interest regarding their potential probiotic applications.

**Keywords:** adaptation; bacteria; growth phases; culturability; long-term survival; viable but nonculturable; *Propionibacterium freudenreichii*.

## Introduction

*Propionibacterium freudenreichii* is an Actinobacterium that is responsible for the opening and aroma of Swiss-type cheeses. PF is also used as a source of probiotics because PF has bifidogenic properties (Isawa *et al.*, 2002), PF is resistant to digestive stress (Jan *et al.*, 2000), and some strains are endowed with anti-inflammatory capabilities that could be used to prevent inflammatory bowel diseases (Okada *et al.*, 2013; Le Maréchal *et al.*, 2015). Its metabolic activity is characterized by its ability to ferment various carbohydrates, the main product being propionic acid (Houwen *et al.*, 1991; Loux *et al.*, 2015). PF is reportedly a hardy bacterial species, which can stay alive and metabolically active for long periods under stressful conditions (i.e. cave-aged Swiss-type cheeses at cool temperatures) (Dalmasso *et al.*, 2012). The genome sequence of the type-strain CIRM-BIA1 revealed the genetic basis for its hardiness (Falentin *et al.*, 2010a). In the laboratory, PF can be stored in a spent medium at room temperature for more than 6 months and remains culturable if inoculated in a fresh medium. Actinobacteria are known to be hardy and extremely resistant to nutrient shortage (Greenblatt *et al.*, 2004). However, except for the *Mycobacterium* genus which has been extensively explored (Ghodbane 2014), little is known about the mechanisms underlying the long term survival of this class.

During their growth and survival, Gram-positive and Gram-negative bacteria pass through different physiological phases (Kolter *et al.*, 1993; Hicks, 2005; Finkel, 2006; Bruno & Freitag, 2011; Wen *et al.*, 2013). In some bacterial species, conventional



exponential and stationary phases are followed successively by the death phase and a long-term survival phase (Finkel, 2006; Helmus *et al.*, 2012). Long-term survival (LTS) can be defined as a phase during which the bacteria develop mechanisms to adapt and remain viable, even after a long period of nutritional shortage (Finkel, 2006; Meena & Rajni, 2010; Bharati *et al.*, 2012; Justice *et al.*, 2014): (i) sporulation (Fimlaid and Shen, 2015), (ii) quiescence/dormancy or (iii) programmed cell death (Finkel, 2006; Rittershaus *et al.*, 2013). During sporulation, a vegetative bacterial cell differentiates into a metabolically dormant cell-type known as a spore, which can reawaken when favourable conditions return (Fimlaid and Shen, 2015). Quiescence is characterized by decelerated or arrested growth in a viable non-replicating state (Rittershaus *et al.*, 2013), which can evolve to a viable but nonculturable (VBNC) state (Ramamurthy *et al.*, 2014). In programmed cell death (bacterial apoptosis), the death and lysis of part of a bacterial population may be profitable to surviving cells, which can catabolize the detritus. Cell death and lysis can be observed by monitoring the fall in the viable count using standard plating assays associated with a leakage of intracellular content that can be assessed biochemically (lactate dehydrogenase quantification) (Wilkinson *et al.*, 1994) or using qPCR methods (Falentin *et al.*, 2010b).

PF is a non-sporulating species, but PF is as yet unknown whether it survives by adopting a strategy of quiescence/dormancy or lysis of part of the bacterial population. During this study, we investigated the LTS (OD, culturability, membrane permeability and lysis) of eight *P. freudenreichii* strains throughout an 11-day period of culture without the addition of any nutrients, in order to investigate inter-strain variability in long-term survival capacities and to better understand the strategies used to ensure survival.

## Materials and Methods

### Bacterial strains

Eight *P. freudenreichii* strains (Table 1) were used during this work. The strains were supplied by the International Centre for Microbial Resources - Bacteria of Food Interest – (Centre International de Ressources Microbiennes - Bactéries d'Intérêt Alimentaire; INRA Rennes).

**Table 1.** *Propionibacterium freudenreichii* strains used during this work.

Strains of <i>P. freudenreichii</i> <sup>a</sup>	Source
CIRM-BIA 1	Reference strain
CIRM-BIA 9	Emmental cheese 1989 Van Niel
CIRM-BIA 119	Gruyère cheese 1973
CIRM-BIA 121	Swiss cheese 1937
CIRM-BIA 127	Emmental cheese 1992
CIRM-BIA 129 alias ITG-P20 <sup>b</sup>	Emmental cheese 1992
CIRM-BIA 138 alias ITG-P9 <sup>b</sup>	Emmental cheese 1992
CIRM-BIA 139 alias ITG-P23 <sup>b</sup>	Emmental cheese 1992

<sup>a</sup> Bacterial strains obtained from the Centre International de Ressources Microbiennes (CIRM) dedicated to Bacteria of Food Interest (Bactéries d'Intérêt Alimentaire, BIA) - (CIRM-BIA, UMR1253, INRA Rennes, France).

<sup>b</sup> Strain CIRM-BIA (alias ITG-P20), strain CIRM-BIA 138 (alias ITG-P9) and strain CIRM-BIA 139 (alias ITG-P23) were kindly supplied by Actalia (Rennes, France).

### Conditions for bacterial growth

All the strains were routinely cultured in YEL medium (pH=7.0) (Malik *et al.*, 1968) with *no agitation* and at 30°C. Each strain was grown in a different test tube for each manipulation, thus preventing the interference of oxygen in the environment and maintaining the *P. freudenreichii* strains under microaerophilic conditions. At T0, an inoculum of 7 log<sub>10</sub> CFU/mL was added to all the tubes.

Growth kinetics were followed by Optical Density measurements using a Model DU 640 spectrophotometer (Beckman Coulter, Fullerton, California, USA) at 650 nm ( $OD_{650}$ ) and by counting the number of CFU (CFU/mL) using the micromethod as described previously (Baron *et al.*, 2006). For enumerations, all the *P. freudenreichii* strains were plated in YEL agar culture medium at 30°C, under microaerophilic conditions, until the visualization of colonies (6 days). In all cases, two technical replicates and two biological replicates were performed.

### **Epifluorescence microscopy**

The eight strains were monitored for 11 days using the LIVE/DEAD® BacLight™ Bacterial Viability Kit – Molecular Probes® (*Life Technologies, USA*), as recommended by the supplier. Microscopic observations were performed at each of the time points used for the  $OD_{650}$  and CFU/mL measurements. Briefly, SYTO9 (which stains all cells green, regardless of their membrane integrity) and propidium iodide (which stains permeabilized cells red) were mixed at a ratio of 1: 1 (5µL SYTO 9 and 5 µL propidium iodide). An 0.5µL aliquot of the mix was added to 100µL of the bacterial culture, vortexed, and incubated *in the dark* for 15 minutes. The bacteria were then analysed using epifluorescence microscopy with a phase contrast Nikon G11 microscope (Nikon D300) and Camera Control Pro 2 software (Nikon, Japan). For each sample, six photos (three photos with a x40 objective and three with a x100 objective) were analysed so as to visually estimate the percentage of permeabilized cells (red / red + green cells) at each time point.

### **DNA extraction of non-lysed cells**

The DNA extraction method only enables the recovery of DNA from non-lysed cells. Samples were collected at three time points (79h, 144h and 168h post-inoculation) during the growth of *P. freudenreichii* CIRM-BIA 138, as described previously (Parayre *et al.*, 2007). Briefly, cells were harvested from 1 mL of the culture sample by centrifugation (5000g, 10 min.). The cell pellets were re-suspended in lysis buffer (Qiagen, Courtaboeuf, France) supplemented with lysozyme (Promega, France; 8mg/sample) and mutanolysin (Sigma, France; 20U/sample). These samples were incubated for 30 min. at 37°C for 1 h, before being treated with proteinase K (0.5mg/sample) prior to purification using the DNeasy Tissue Kit, as recommended by the supplier (Qiagen, Courtaboeuf, France). Three biological replicates and two technical replicates were used for each time point.

### **Quantification of bacterial cells using quantitative PCR**

To determine any cell lysis of CIRM-BIA 138 after entry into stationary phase and just before entering the LTS phase, quantitative PCR (qPCR) was used to quantify the non-lysed bacterial cells throughout growth, according to the method described by Falentin *et al.* (2010b). Briefly, the qPCR targeted *tuf*, a gene present in a single copy in the *P. freudenreichii* chromosome, using primers designed on the *tuf* sequence (F: CGAACGAGTTCCACTGCGGGT; R: GCAACATCGGCACCATCGGAC). Amplifications were performed exactly as described previously. Cycle thresholds (Ct) were determined after manual adjustment of the threshold within the linear part of qPCR logarithmic amplification curves. CIRM-BIA 138 samples were collected at 79, 144, and 168 hours for qPCR determination. Each experiment was carried out in triplicate (independent biological replicates). We checked by a spiking experiment that

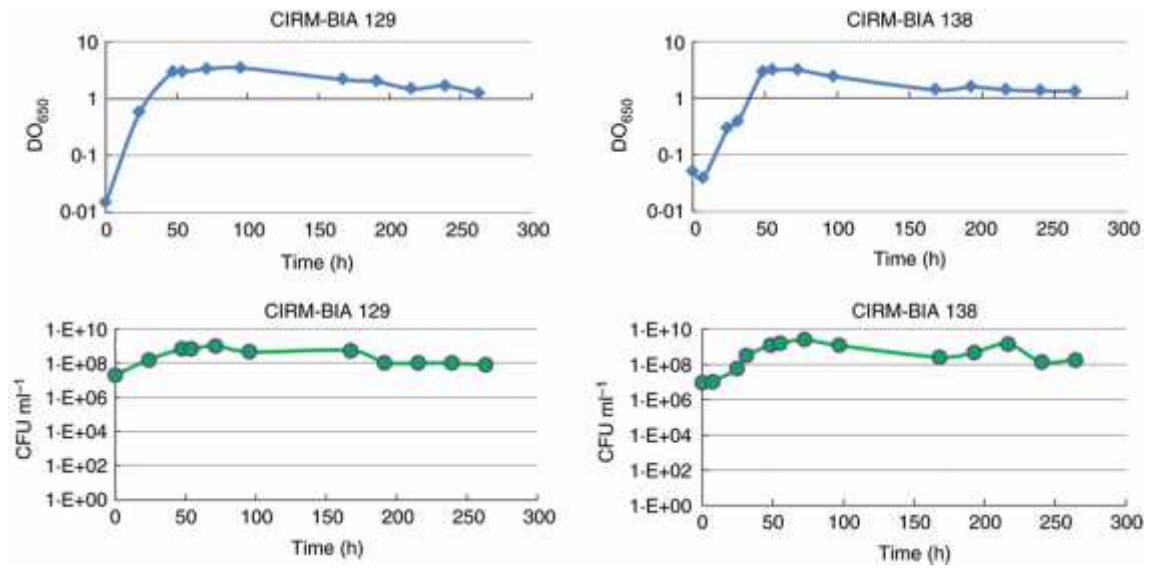
lyzed cells could not contaminate qPCR results. A sample of entire cells ( $8 \log_{10}$  CFU) was spiked with ( $9 \log_{10}$  CFU) lyzed cells (the lysis was performed as described below). Extraction and qPCR on entire cells and spiked sample gave the same Ct. The qPCR results will give a quantification of entire cells only.

## Results

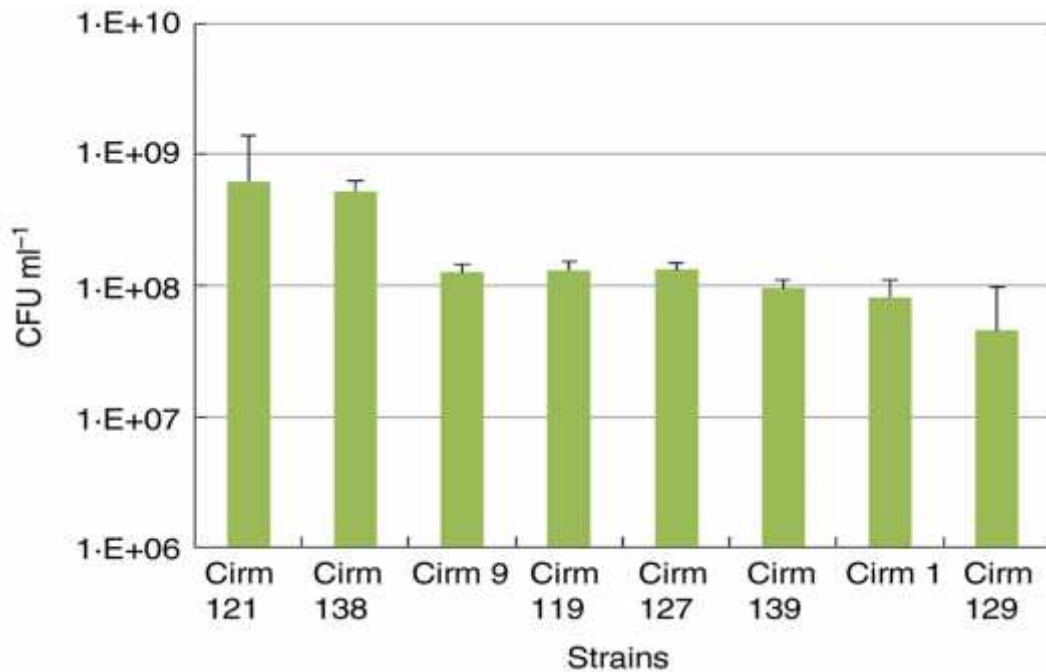
### **Growth and survival rate of *Propionibacterium freudenreichii* under laboratory conditions.**

The growth of the eight *P. freudenreichii* strains was monitored spectrophotometrically by measuring the optical density at 650 nm ( $OD_{650}$ ), and by plate counting (CFU/mL) throughout an 11 day period. Strains were inoculated at  $7 \log_{10}$  CFU/mL in fresh YEL medium. Both of the monitoring methods showed that *P. freudenreichii* growth curves had roughly similar profiles despite variations in terms of their final culturability at 11 days. The curves established over a prolonged incubation revealed several phases, as shown for CIRM-BIA129 and CIRM-BIA138 (Figure 1. See all growth curves in the supplemental data, Figure S1). Under the conditions applied, the cultures entered the exponential growth phase to reach a maximum  $OD_{650}$  of 3 to 4 (for CIRM-BIA 138) and a population of 9 to  $9.6 \log_{10}$  CFU/mL between 50 h and 72 h. The  $OD_{650}$  curves then decreased steadily for about 100h, which corresponded to entry into the stationary phase, and reached a plateau of culturability at around 168 h, for a population of 8 to  $8.7 \log_{10}$  CFU/mL, depending on the strain (Figure 1) but always with a lower  $OD_{650}$ . The latter phase was designated the long term survival (LTS) phase. The other six *P. freudenreichii* growth curves were similar (Supplementary data Figure S1). In terms of culturability, all the strains survived well with a population of around  $8 \log_{10}$  CFU/mL after 11 days in the spent medium, i.e. without the addition of any nutrients (Figure 2, Table S1). Interestingly, two strains, CIRM-BIA121 and CIRM-BIA138, survived better than the six others (Figure 2) and maintained an average population of  $8.2 \log_{10}$  CFU/ mL at day 11, suggesting a strain-dependency in ability of *P. freudenreichii* for long-term survival. On Petri plates, colonies of *P. freudenreichii* maintained the same size and aspect whereas photon microscopy images revealed

bacteria of the same size (not shown) but single cells in the exponential phase and groups of two to five cells in the stationary phase.



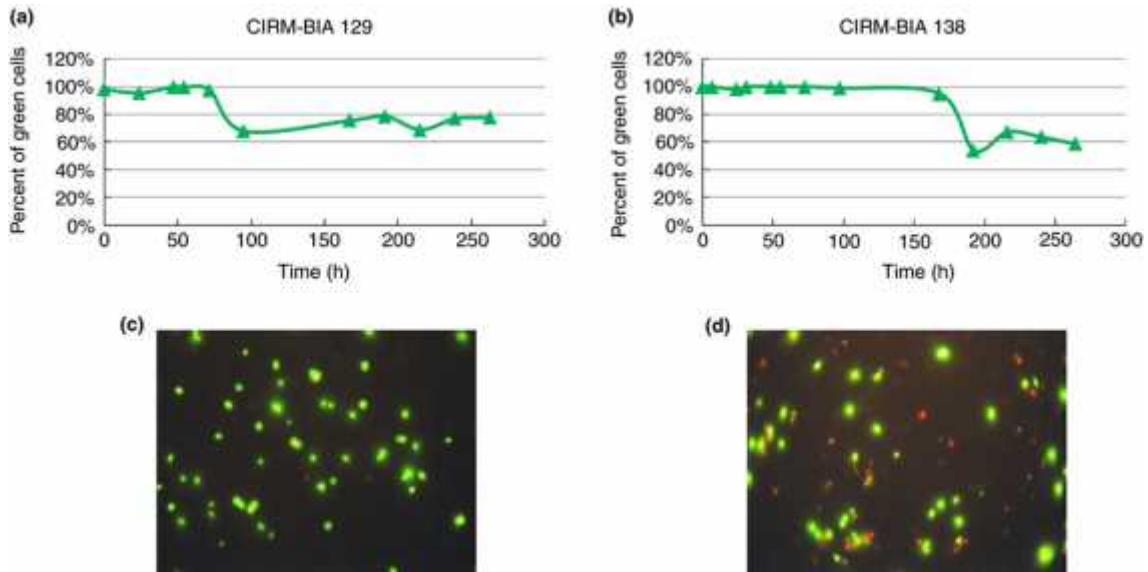
**Figure 1:** Growth curves of *P. freudenreichii* strains CIRM-BIA129 and CIRM-BIA138 determined at OD650 (upper panels), and plate counting (CFU ml<sup>-1</sup>, lower panels).



**Figure 2:** Viability of the eight *P. freudenreichii* strains (in CFU ml<sup>-1</sup>) after 11 days in the spent culture medium (YEL) based on a mean of two technical replicates and two biological replicates. Upper bars are standard deviations.

## Evaluation of cell envelope permeability

LIVE/DEAD staining assays showed that the numbers of intact cells (stained green by SYTO9) remained stable (100% index) for each of the eight strains throughout the exponential and early stationary phases (Table S2). The numbers of intact cells then started decreasing at 50 or 150h post-inoculation and dropped to 80% and 58.6% for CIRM-BIA 129 and CIRM-BIA 138, respectively, 11 days post-inoculation. The numbers of cells with a non-permeabilized membrane (stained in green) decreased from 50 h to 150 h onwards, depending on the strains, to reach a minimum proportion of 50% at day 11 (Figure 3), suggesting that the number of dying or dead (but non-lysed) cells increased during the stationary and LTS phases.



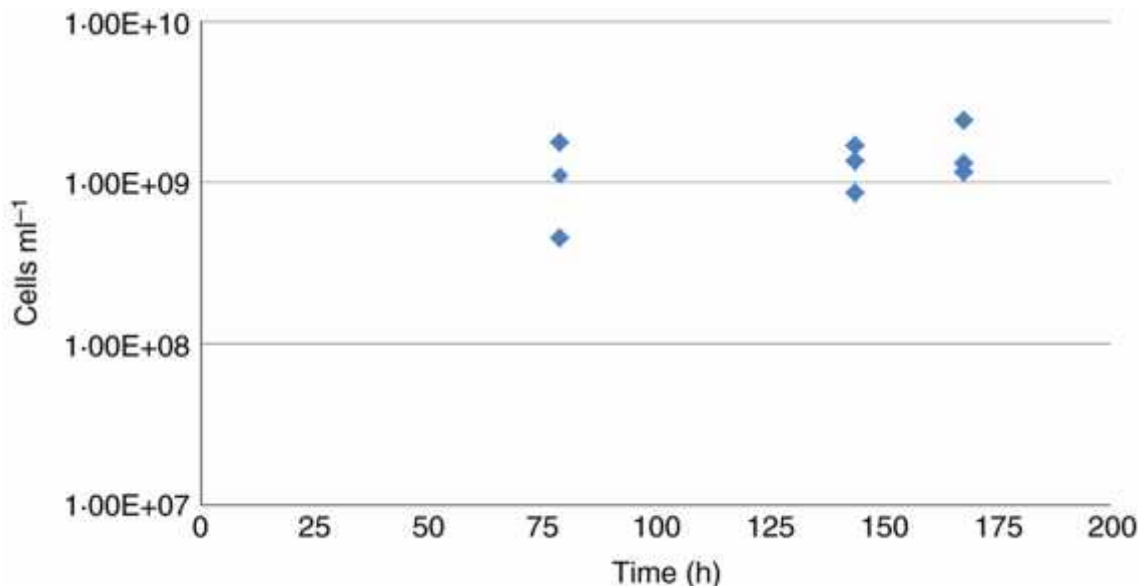
**Figure 3:** (A) Proportion of non-permeabilized cells during the growth and survival of *P. freudenreichii* strains CIRM-BIA129 (a) and 138 (b) measured by LIVE/DEAD assay followed by epifluorescence microscopy (mean values of three epifluorescence microscopy fields at x40 magnification, and three fields at x100 magnification). Representative epifluorescence microscopy views of *P. freudenreichii* (CIRM-BIA 138)



cultures after LIVE/DEAD staining during exponential growth (c) and after 11 days of growth (d).

### Assessment of lysis by quantifying bacterial cells using qPCR

Because both plate counting assays and epifluorescence microscopy showed a reduction in culturability and an increase in membrane permeabilization, quantitative PCR was used to assess bacterial lysis. CIRM-BIA 138 (the strain with the highest population at day 11) cells were pelleted at three different time points before and after the observed loss of culturability. Quantitative PCR targeting the *tuf* gene was carried out on the DNA extracted from the pelleted cells. Under these conditions, lysed cells cannot normally be recovered (Falentin *et al.*, 2012). The copy number of the *tuf* gene did not vary significantly in the CIRM-BIA138 culture, demonstrating that the bacterial cells did not undergo lysis when entering the LTS phase (Figure 4).



**Figure 4:** Quantification of bacterial cells during the growth of *P. freudenreichii* CIRM-BIA138. Each point gives the copy number of *tuf* gene determined by qPCR on DNA extracted from *P. freudenreichii* CIRM-BIA138 (three biological replicates).

## Discussion

### ***P. freudenreichii* has a high survival rate at day 11**

*P. freudenreichii* belongs to the Actinobacteria, a group of Gram positive bacteria with high GC content genomes. This group is known for its considerable versatility (Gao & Gupta, 2012), and includes members that can survive under extreme conditions (Greenblatt *et al.*, 2004). *P. freudenreichii* is reported to be a hardy species widely used in the cheese-making industry and is also well-documented for its probiotic potential (Jan *et al.*, 2000). Genome sequence analysis has revealed the genetic basis for its hardiness: (i) its ability to store polyphosphates as an energy reserve and to utilize them instead of ATP, (ii) its ability to store glycogen as a carbon source, and (iii) several copies of chaperones to cope with different stresses (Falentin *et al.* 2010a). The eight *P. freudenreichii* strains studied here survived well during a period of 11 days under nutritional shortage, according to the optical density and plate counting assays performed (around  $8 \log_{10}$  CFU/mL). In YEL medium, lactate - the preferential carbon source for PF- was exhausted two days after inoculation (Dalmasso *et al.* 2012). LIVE/DEAD staining followed by epifluorescence microscopy ascertained that more than 50% of the *P. freudenreichii* populations had an intact (non-permeabilized) membrane at this date. The growth curves for CIRM-BIA 129 and 138 (Fig. 1) displayed some fluctuations during long-term incubation. We suggest that intermittent cryptic regrowth might be responsible for this variability. These cryptic growths may result from some specific metabolisms induced to adapt to starvation or the growth of a sub-population of mutants derived from the inoculated strain and with a growth advantage that was explained in the review by Finkel *et al.* (2006).

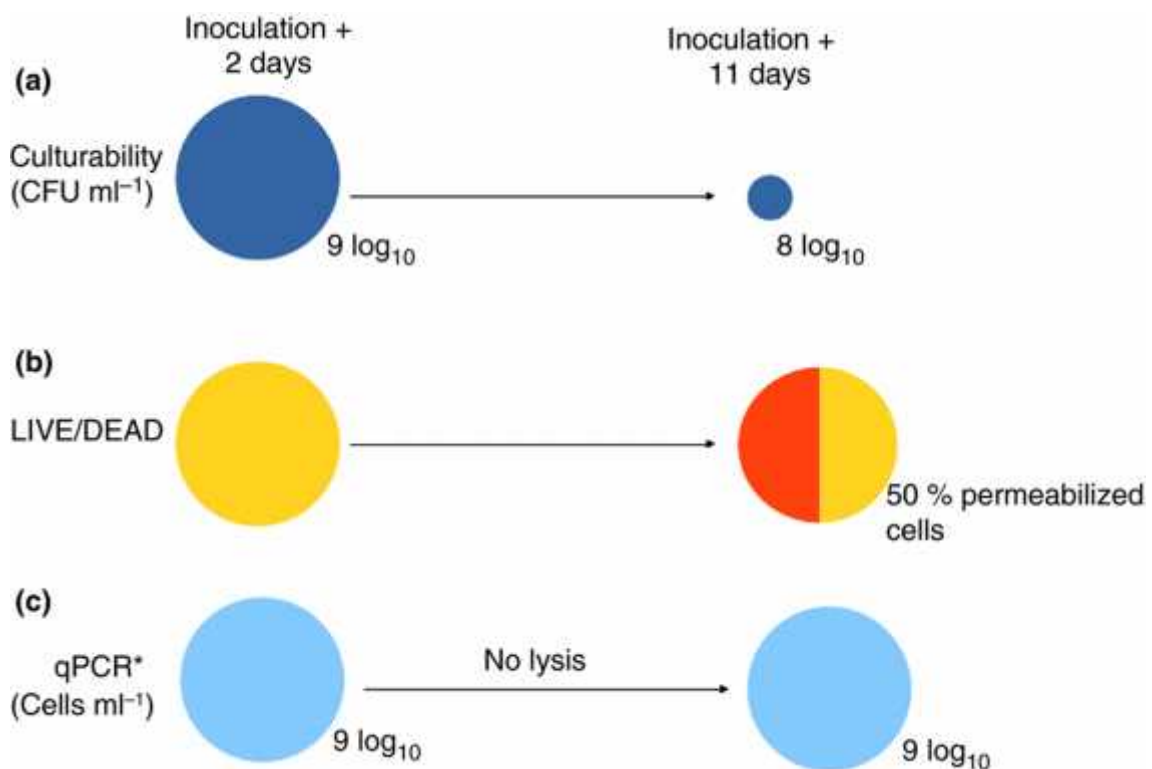
### **The best surviving strain did not use partial lysis to sustain viability**

Previous studies had shown that some lactic acid bacteria present in dairy products can lyse under difficult conditions and that the cell debris released into the culture medium may serve as a nutrient source to ensure survival of the non-lysed sub-population (Lortal *et al.*, 1997; Deutsch *et al.*, 2002; Lortal & Chapot-Chartier, 2005). A study of the autolysis of *Lactobacillus* and dairy propionic bacteria showed that this autolysis is associated with several factors, including among others the type of strain, autolysins, pH, temperature and ionic strength (Lortal *et al.*, 1997). For CIRM-BIA 138, the quantification of intact cells using quantitative PCR showed that under the study conditions applied here, lysis could be ruled out. Whether this is a general feature of *P. freudenreichii* still needs to be determined. The absence of lysis (i) could be confirmed by quantitative PCR on cells exposed to propidium monoazide, so as to prevent any contamination by DNA from dead cells, (ii) needs to be confirmed in other *P. freudenreichii* strains and under different conditions before it is generally accepted.

### ***P. freudenreichii* probably enters a viable but nonculturable (VBNC) state**

The experimental data reported here offer a range of presumptions suggesting that *P. freudenreichii* enters into a VBNC state. VBNC is a survival strategy adopted by numerous bacteria in response to adverse conditions such as starvation, extreme temperature, pH, oxygen stress, antibiotic pressure, etc. Once in a VBNC state, the bacteria no longer grow in routine laboratory medium and conditions, but they nevertheless remain viable and can be resuscitated when more favourable conditions are provided (Ramamurthy *et al.*, 2014). The methods generally used to detect VBNC bacteria include bright field microscopy with nalidixic acid; fluorescent microscopy,

and especially use of the BacLight LIVE/DEAD kit; and molecular techniques (DNA or RNA) (Oliver, 2009; Fakruddin *et al.*, 2013). All of the eight *P. freudenreichii* strains studied here displayed a loss of culturability of  $\sim 1 \log_{10}$  CFU/mL (Table S1 and Figure 5), when this parameter was compared between 50-72 hours and 11 days post-inoculation. LIVE/DEAD staining showed that about half of the bacterial cells had a permeabilized membrane. Using quantitative PCR, an absence of lysis was shown on strain CIRM-BIA138. These findings suggest that at least a part of the *P. freudenreichii* population entered a VBNC state and that the entry into LTS is concomitant with entry into a VBNC state, at least for  $1 \log_{10}$  (i.e. 90%) CFU/mL. Since the total bacterial population was stained either green (50%) or red (50%), we suggest that VBNC cells (90%) can take up both colours and that LIVE/DEAD staining cannot be used to quantify VBNC cells.



**Figure 5:** Fate of CIRM-BIA138 bacterial cells between 50-72 h post-inoculation and the long-term survival phase 11 days post-inoculation assessed by (a) culturability on

Petri plates, (b) LIVE/DEAD staining followed by epifluorescence and (c) quantification of intact cells by quantitative PCR.

The VBNC state has been described in various Gram positive bacteria, such as *Enterococcus faecalis*, *Enterococcus hirae*, *Listeria monocytogenes*, and *Mycobacterium tuberculosis* (Oliver et al., 2009), and in Gram negative bacteria such as *Vibrio cholerae* (Gonzales-Escalona et al. 2006), *Escherichia coli* (Darcan et al., 2009) and *Helicobacter pylori* (Casasola-Rodríguez et al., 2013). Many of the studies on the VBNC state have focused on pathogenic bacteria (Ramamurthy et al. 2014). This state has also recently been reported for bacteria of food interest and particularly in ripened cheese, suggesting a role for these bacteria in the organoleptic qualities of cheeses (Martin-Platero et al. 2008, Desfossés-Foucault et al. 2012, Ruggirello et al. 2014).

In conclusion, our findings suggest that one strategy for bacterial survival is dormancy, rather than partial bacterial lysis, in order to sustain the remaining cells. During the past decade, many studies have investigated the bacterial mechanisms underlying LTS in various fields: environmental, industrial, medical and biotechnological microbiology. This study represents the first work involving the phenotypic characterization of *P. freudenreichii* over an 11-day period without the addition of nutrients, and revealed different phases of growth, membrane permeabilization and entry into dormancy and a VBNC state to ensure LTS. A VBNC state may however have an important impact on the technological uses of these bacteria, as either starters in cheese-making or probiotics, since viability and metabolic activity are important traits in their wide range of potential uses. This therefore opens the way for further investigations on a molecular basis of the LTS of *P. freudenreichii*. The availability of genome sequences (Loux et al., 2015) will greatly assist such investigations using comparative genomics and transcriptomics approaches in order to

understand the organization, genes and metabolic pathways required for LTS. Assessing viability and understanding the mechanisms underlying the survival of *P. freudenreichii* should be of considerable interest regarding the development of effective bacterial probiotics.

### **Acknowledgements**

This work was supported by the CAPES-COFECUB French-Brazilian Cooperation Programme. We also wish to thank the Institut National de la Recherche Agronomique (INRA, UMR 1253, STLO, Rennes, France) for its financial and practical support. Our thanks also go to Professor Mary Bret and Vicky Hawken for English revision of the paper.

### **Conflict of Interest section**

The authors have no conflict of interests to declare.

## References

- Baron, F., Cochet, M.-F., Ablain, W., Grosset, N., Madec, M.-N., Gonnet, F., Jan, S. and Gautier, M. (2006) Rapid and cost-effective method for micro-organism enumeration based on miniaturization of the conventional plate-counting technique. *Le Lait* **86**, 251–257.
- Bharati, B.K., Sharma, I.M., Kasetty, S., Kumar, M., Mukherjee, R. and Chatterji, D. (2012) A full-length bifunctional protein involved in c-di-GMP turnover is required for long-term survival under nutrient starvation in *Mycobacterium smegmatis*. *Microbiology* **158**,1415–1427.
- Brandhorst, D., Brandhorst, H., Hering, B.J. and Bretzel, R.G. (1999) Long-term survival, morphology and in vitro function of isolated pig islets under different culture conditions. *Transplantation* **67**,1533–1541.
- Bruno, J.C. and Freitag, N.E. (2011) *Listeria monocytogenes* adapts to long-term stationary phase survival without compromising bacterial virulence. *FEMS Microbiol Lett* **323**,171–179.
- Casasola-Rodríguez, B., Orta de Velásquez, M.T., Luqueño-Martínez, V.G. and Monje-Ramírez, I. (2013) Quantification of *Helicobacter pylori* in the viable but nonculturable state by quantitative PCR in water disinfected with ozone. *Water Sci Technol J Int Assoc Water Pollut Res* **68**,2468–2472.
- Dalmasso, M., Aubert, J., Briard-Bion, V., Chuat V., Deutsch S.-M., Even, S., Falentin, H., Jan, G., Jardin, J., Maillard, M.-B., Parayre, S., Piot, M., Tanskanen, J., Thierry, A. (2012) A Temporal -omic Study of Propionibacterium freudenreichii CIRM-BIA1T Adaptation Strategies in Conditions Mimicking Cheese Ripening in the Cold. *PLoS ONE* **7**, e29083.
- Darcan, C., Ozkanca, R., Idil, O. and Flint, K.P. (2009) Viable but non-culturable state (VBNC) of *Escherichia coli* related to EnvZ under the effect of pH, starvation and osmotic stress in sea water. *Pol J Microbiol Pol Tow Mikrobiol Pol Soc Microbiol* **58**,307–317.
- Desfossés-Foucault, É., Dussault-Lepage, V., Le Boucher, C., Savard, P., LaPointe, G., Roy, D. (2012) Assessment of probiotic viability during Cheddar cheese manufacture and ripening using propidium monoazide-PCR quantification. *Front Microbiol*, **3**:1-11.
- Deutsch, S.-M., Ferain, T., Delcour, J. and Lortal, S. (2002) Lysis of lysogenic strains of *Lactobacillus helveticus* in Swiss cheeses and first evidence of concomitant *Streptococcus thermophilus* lysis. *Int Dairy J* **12**,591–600.
- Fakruddin, M., Mannan, K.S.B. and Andrews, S. (2013) Viable but Non Culturable Bacteria: Food Safety and Public Health Perspective. *ISRN Microbiol* **2013**,1–6.
- Falentin, H., Deutsch, S.-M., Jan, G., Loux, V., Thierry, A., Parayre, S., Maillard, M.-B., Dherbécourt, J., Cousin, F.J., Jardin, J., Siguier, P., Couloux, A., Barbe, V., Vacherie, B., Wincker, P., Gibrat, J.-F., Gaillardin, C., Lortal, S. (2010a) The

Complete Genome of *Propionibacterium freudenreichii* CIRM-BIA1T, a Hardy Actinobacterium with Food and Probiotic Applications. *PLoS ONE* **5**:e11748.

- Falentin, H., Postollec, F., Parayre, S., Henaff, N., Le Bivic, P., Richoux, R., Thierry, A. and Sohier, D. (2010b) Specific metabolic activity of ripening bacteria quantified by real-time reverse transcription PCR throughout Emmental cheese manufacture. *Int J Food Microbiol* **144**,10–19.
- Falentin, H., Henaff, N., Le Bivic, P., Deutsch, S.-M., Parayre, S., Richoux, R., Sohier, D., Thierry, A., Lortal, S. and Postollec, F. (2012) Reverse transcription quantitative PCR revealed persistency of thermophilic lactic acid bacteria metabolic activity until the end of the ripening of Emmental cheese. *Food Microbiol* **29**,132–140.
- Fimlaid, K.A., Shen, A.(2015) Diverse mechanisms regulate sporulation sigma factor activity in the Firmicutes. *Curr Opin Microbiol* **24**,88–95.
- Finkel, S.E. (2006) Long-term survival during stationary phase: evolution and the GASP phenotype. *Nat Rev Microbiol* **4**,113–120.
- Gao, B. and Gupta, R.S.(2012) Phylogenetic Framework and Molecular Signatures for the Main Clades of the Phylum Actinobacteria. *Microbiol Mol Biol Rev* **76**,66–112.
- Ghodbane, R., Medie, F.M., Lepidi, H., Nappez, C., Drancourt, M. (2014) Long-term survival of tuberculosis complex mycobacteria in soil. *Microbiology* **160**,496–501.
- Gonzalez-Escalona, N., Fey, A., Hofle, M.G., Espejo, R.T., Guzman, C.A. (2006) Quantitative reverse transcription polymerase chain reaction analysis of *Vibrio cholerae* cells entering the viable but non-culturable state and starvation in response to cold shock. *Environ Microbiol*, **8**:658–666.
- Greenblatt, C.L., Baum, J., Klein, B.Y., Nachshon, S., Koltunov, V. and Cano, R.J. (2004) *Micrococcus luteus* - Survival in Amber. *Microb Ecol* **48**,120–127.
- Helmus, R.A., Liermann, L.J., Brantley, S.L. and Tien, M. (2012) Growth advantage in stationary-phase (GASP) phenotype in long-term survival strains of *Geobacter sulfurreducens*. *FEMS Microbiol Ecol* **79**,218–228.
- Hicks, W.M. (2005) Recovery from long-term stationary phase and stress survival in *Escherichia coli* require the L-isoaspartyl protein carboxyl methyltransferase at alkaline pH. *Microbiology* **151**,2151–2158.
- Houwen, F.P., Dijkema, C., Stams, A.J.M. and Zehnder, A.J.B (1991) Propionate metabolism in anaerobic bacteria; determination of carboxylation reactions with <sup>13</sup>C-NMR spectroscopy. *Biochim Biophys Acta BBA - Bioenerg* **1056**,126–132.
- Isawa, K., Hojo, K., Yoda, N., Kamiyama, T., Makino, S., Saito, M., Sugano, H., Mizoguchi, C., Kurama, S., Shibasaki, M., Endo, N., Sato, Y. (2002) Isolation and Identification of a New Bifidogenic Growth Stimulator Produced by *Propionibacterium freudenreichii* ET-3. *Biosci Biotechnol Biochem* **66**,679–681.



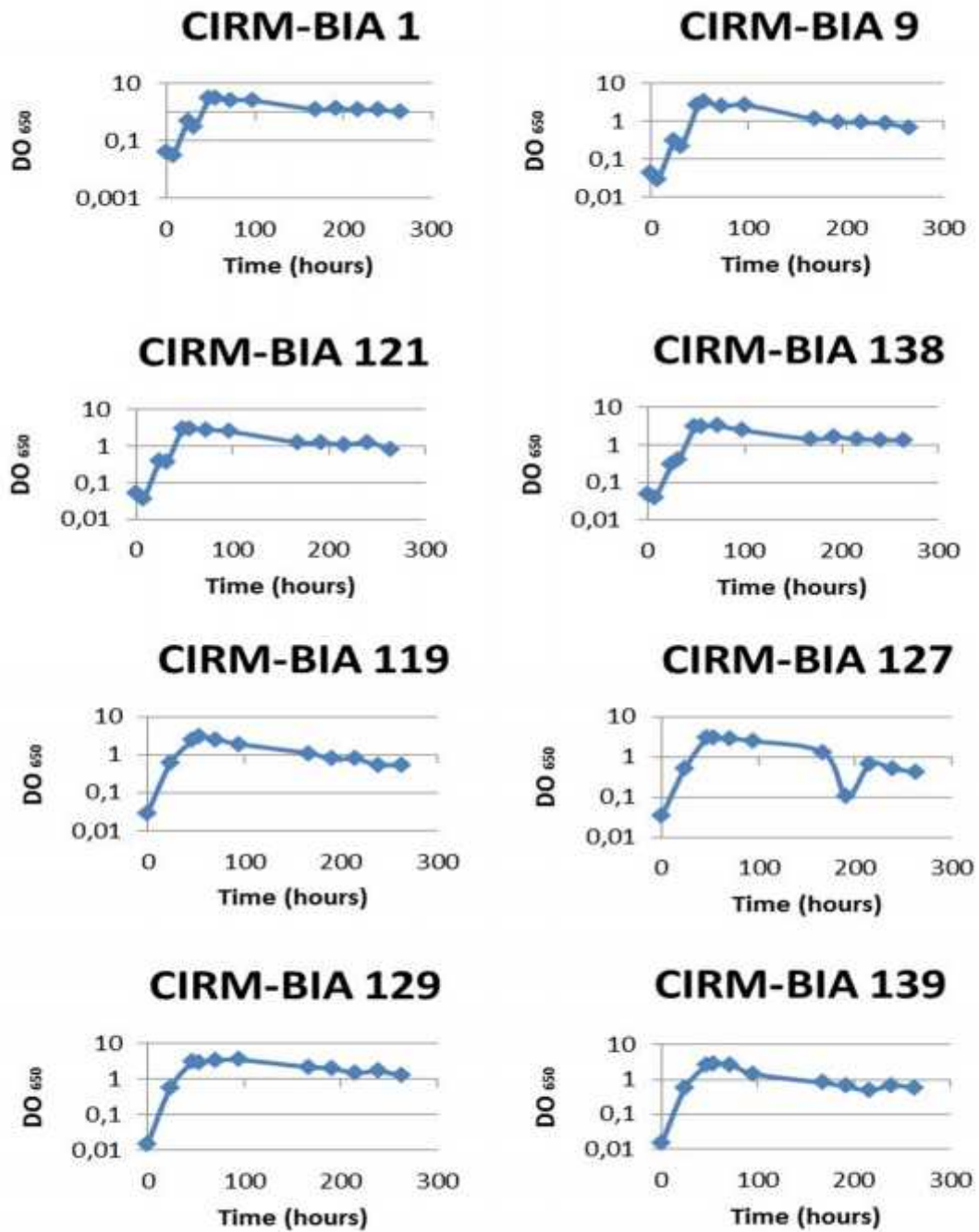
- Jan, G., Rouault, A. and Maubois, J.-L. (2000) Acid stress susceptibility and acid adaptation of *Propionibacterium freudenreichii* subsp. *shermanii*. *Le Lait* **80**,325–336.
- Justice, S.S., Harrison, A., Becknell, B. and Mason, K.M. (2014) Bacterial differentiation, development, and disease: mechanisms for survival. *FEMS Microbiol Lett* **360**,1–8.
- Kolter, R., Siegele, D.A. and Tormo, A. (1993) The Stationary Phase of The Bacterial Life Cycle. *Annu Rev Microbiol* **47**,855–874.
- Le Maréchal, C., Peton, V., Plé, C., Vroland, C., Jardin, J., Briard-Bion, V., Durant, G., Chuat, V., Loux, V., Foligné, B., Deutsch, S.-M., Falentin, H., Jan, G.(2015) Surface proteins of *Propionibacterium freudenreichii* are involved in its anti-inflammatory properties. *J Proteomics* **113**,447–461.
- Lortal, S. and Chapot-Chartier, M.-P. (2005) Role, mechanisms and control of lactic acid bacteria lysis in cheese. *Int Dairy J* **15**,857–871.
- Lortal, S., Lemée, R. and Valence, F. (1997) Autolysis of thermophilic lactobacilli and dairy propionibacteria: a review. *Le Lait* **77**,133–150.
- Loux, L., Mariadassou, M., Almeida, S., Chiapello, H., Hammani, A., Buratti, J. J., Gendrault, A., Barbe, V., Aury, J.-M., Deutsch, S.-M., Parayre, S., Madec, M.-N., Chuat, V., Jan, G., Peterlongo, P., Azevedo, V., Le Loir, Y., Falentin, H. (2015) Mutations and genomic islands explain the strain dependancy of sugar utilization among 21 strains of *Propionibacterium freudenreichii*. *BMC Genomics* **16**,296.
- Malik, A.C., Reinbold, G.W. and Vedamuthu, E.R. (1968) An evaluation of the taxonomy of *Propionibacterium*. *Can J Microbiol* **14**,1185–1191.
- Martin-Platero, A.M., Valdivia, E., Maqueda, M., Martin-Sanchez, I., Martinez-Bueno, M. (2008) Polyphasic approach to bacterial dynamics during the ripening of Spanish farmhouse cheese, using culture-dependent and -independent methods. *Appl Environ Microbiol*, **74**:5662–5673.
- Meena, L.S., and Rajni, (2010) Survival mechanisms of pathogenic *Mycobacterium tuberculosis* H37Rv. *FEBS J* **277**,2416–2427.
- Okada, Y., Tsuzuki, Y., Narimatsu, K., Sato, H., Ueda, T., Hozumi, H., Sato, S., Hokari, R., Kurihara, C., Komoto, S., Watanabe, C., Tomita, K., Kawaguchi, A., Nagao, S., Miura, S.(2013) 1,4-Dihydroxy-2-naphthoic acid from *Propionibacterium freudenreichii* reduces inflammation in interleukin-10-deficient mice with colitis by suppressing macrophage-derived proinflammatory cytokines. *J Leukoc Biol* **94**,473–480.
- Oliver, J.D. (2009) Recent findings on the viable but nonculturable state in pathogenic bacteria. *FEMS Microbiol Rev.* **34**,415-25.
- Parayre, S., Falentin, H., Madec, M.-N., Sivieri, K., Le Dizes, A.-S., Sohier, D. and Lortal, S. (2007) Easy DNA extraction method and optimisation of PCR-

- Temporal Temperature Gel Electrophoresis to identify the predominant high and low GC-content bacteria from dairy products. *J Microbiol Methods* **69**,431–441.
- Ramamurthy, T., Ghosh, A., Pazhani, G.P. and Shinoda, S. (2014) Current Perspectives on Viable but Non-Culturable (VBNC) Pathogenic Bacteria. *Front Public Health* **2**.
- Rittershaus, E.S.C., Baek, S.-H. and Sasseti, C.M. (2013) The Normalcy of Dormancy: Common Themes in Microbial Quiescence. *Cell Host Microbe* **13**,643–651.
- Ruggirello, M., Dolci, P., Cocolin, L. (2014) Detection and viability of *Lactococcus lactis* throughout cheese ripening. *PLoS ONE* 2014, **9**:e114280.
- Wen, J., Karthikeyan, S., Hawkins, J., Anantheswaran, R.C. and Knabel, S.J.(2013) *Listeria monocytogenes* responds to cell density as it transitions to the long-term-survival phase. *Int J Food Microbiol* **165**,326–331.
- Wilkinson, M.G., Guinee, T.P. and Fox, P.F. (1994) Factors which may influence the determination of autolysis of starter bacteria during cheddar cheese ripening. *Int Dairy J* **4**,141–160.

## Supporting Information

Additional supporting information can be found in the online version of this article.

**Figure S1.** Kinetics of growth and survival measured by optical density in eight strains of *Propionibacterium freudenreichii*.



**Table S1.** Kinetics of the growth and survival of eight *P. freudenreichii* strains measured by culturability in Petri dishes.

Table S1

Time (h)	Exp 1	Exp 1	Exp 1	Exp 1
	<b>Cirm 1</b>	<b>Cirm 9</b>	<b>Cirm 121</b>	<b>Cirm 138</b>
0,00	1,33E+07	1,86E+07	1,28E+07	9,65E+06
7,00	1,39E+07	1,45E+07	1,19E+07	1,11E+07
24,30	1,49E+08	1,58E+08	1,42E+08	5,60E+07
31,30	1,47E+08	1,42E+08	3,86E+08	3,18E+08
48,30	3,51E+08	1,33E+09	5,76E+08	1,20E+09
55,30	1,43E+09	2,83E+09	1,95E+09	1,55E+09
72,30	1,08E+09	1,18E+09	3,55E+09	2,60E+09
97,00	5,51E+08	1,35E+09	8,26E+08	1,20E+09
168,00	2,45E+08	7,22E+08	2,95E+08	2,59E+08
192,30	1,59E+08	6,39E+08	1,35E+08	4,73E+08
216,30	1,56E+08	4,08E+08	2,62E+08	1,47E+09
240,30	1,34E+08	1,41E+08	1,01E+08	1,40E+08
264,30	1,01E+08	1,14E+08	2,00E+08	1,89E+08
	8,0	8,1	8,3	8,3

-average day 11            1,14E+008    8,058105595  
 -standard error day 11    4,64E+007    7,666202446

**Exp 1**

max day 11            2,00E+008    

<b>Cirm 121</b>
-----------------

  
 min day 11            1,01E+008    

<b>Cimr1</b>
--------------

Time (h)	Exp 2	Exp 2	Exp 2	Exp 2
	<b>Cimr 119</b>	<b>Cirm 127</b>	<b>Cirm 129</b>	<b>Cirm 139</b>
0,00	1,50E+07	2,55E+07	2,05E+07	6,45E+07
24,00	2,80E+08	6,10E+08	1,65E+08	4,43E+08
47,30	7,68E+08	1,09E+09	7,03E+08	1,51E+09
54,30	1,07E+09	1,49E+09	7,00E+08	2,13E+09
71,30	8,11E+08	3,27E+09	1,07E+09	4,35E+09
95,00	8,50E+08	4,25E+09	4,75E+08	2,03E+09
167,30	1,24E+08	6,40E+08	5,60E+08	6,03E+08
191,30	1,08E+08	4,58E+08	1,13E+08	5,78E+08
215,30	1,65E+08	3,00E+08	1,03E+08	2,93E+08
239,00	1,05E+08	1,57E+08	1,04E+08	1,45E+08
263,00	1,15E+08	1,43E+08	8,28E+07	8,53E+07
	8,1	8,2	7,9	7,9

**Exp 2**

max day 11    1,43E+008    

<b>Cirm 127</b>
-----------------

  
 min day 11    8,28E+007    

<b>Cirm 129</b>
-----------------

	Exp 3	Exp 3	Exp 3	Exp 3
Time (h)	<b>Cimr 9</b>	<b>Cimr 121</b>	<b>Cimr 127</b>	<b>Cimr 138</b>
0,00	1,23E+07	2,35E+07	1,63E+07	1,14E+07

24,00	3,88E+08	4,73E+08	3,33E+08	4,83E+08
31,30	8,38E+08	1,87E+09	2,43E+08	3,53E+08
47,00	1,31E+09	2,78E+09	1,09E+09	1,30E+09
53,00	1,27E+09	2,03E+09	7,03E+08	7,05E+08
71,00	1,36E+09	1,03E+09	1,25E+09	2,98E+09
95,00	1,46E+09	1,32E+09	1,75E+09	1,52E+09
168,00	5,75E+08	2,45E+08	1,70E+08	2,33E+08
192,00	4,60E+08	3,35E+08	1,65E+08	1,43E+08
216,00	4,58E+08	4,13E+08	1,50E+08	1,77E+08
240,00	3,95E+08	3,28E+08	1,20E+08	1,16E+08
263,30	1,39E+08	9,23E+07	1,20E+08	1,29E+08
	8,1	8,0	8,1	8,1

### Exp 3

max day 11	1,39E+008	<b>Cimr 9</b>
min day 11	9,23E+007	<b>Cimr 121</b>

	Exp 4	Exp 4	Exp 4	Exp 4
Time (h)	<b>Cimr 1</b>	<b>Cimr 119</b>	<b>Cimr 129</b>	<b>Cimr 139</b>
0,00	1,24E+07	1,48E+07	1,80E+07	1,27E+07

23,00	1,31E+08	1,58E+08	3,13E+07	1,42E+08
29,00	7,00E+08	3,08E+08	1,98E+08	4,50E+08
47,00	1,35E+09	8,05E+08	5,95E+08	1,47E+09
53,00	1,48E+09	1,05E+09	6,68E+08	1,26E+09
71,00	1,03E+09	9,55E+08	9,13E+08	2,80E+09
94,30	4,55E+08	8,70E+08	9,15E+08	9,75E+08
166,00	2,58E+08	7,28E+08	1,73E+08	5,85E+08
189,30	1,13E+08	1,50E+08	7,25E+07	4,18E+08
212,00	6,43E+07	1,19E+08	4,50E+07	4,38E+08
235,30	7,33E+07	1,00E+08	8,83E+07	3,80E+08
259,00	6,15E+07	1,45E+08	7,78E+06	1,05E+08
	7,8	8,2	6,9	8,0

### Exp 4

max day 11	1,45E+008	<b>Cimr 119</b>
min day 11	7,78E+006	<b>Cimr 129</b>

**Table S2.** Proportion of non-permeabilized cells during the growth and survival of eight *P. freudenreichii* strains measured by LIVE/DEAD staining followed by epifluorescence microscopy.

	Exp 1	Exp 1	Exp 1	Exp 1
Time (h)	<b>Cirm1</b>	<b>Cirm 9</b>	<b>Cirm 121</b>	<b>Cirm 138</b>
0,00	100,0%	100,0%	100,0%	100,0%
7,00	100,0%	100,0%	100,0%	100,0%
24,30	98,00%	90,00%	98,00%	98,00%
31,30	100,00%	100,00%	100,00%	100,00%
48,30	100,00%	100,00%	100,00%	100,00%
55,30	100,00%	100,00%	100,00%	100,00%
72,30	100,00%	99,00%	98,00%	100,00%
97,00	99,00%	100,00%	100,00%	99,00%
168,00	85,00%	95,00%	90,00%	95,00%
192,30	40,00%	64,00%	57,50%	54,00%
216,30	54,40%	74,25%	46,00%	67,33%
240,30	37,50%	64,50%	45,66%	63,50%
264,30	51,33%	56,83%	32,40%	58,55%

	Exp 2	Exp 2	Exp 2	Exp 2
Time (h)	<b>Cirm 119</b>	<b>Cirm 127</b>	<b>Cirm 129</b>	<b>Cirm 139</b>
0,00	90,6%	96,3%	98,0%	94,0%
24,00	96,5%	98,6%	95,4%	100,0%
47,30	96,90%	98,80%	100,00%	100,00%
54,30	99,83%	100,00%	99,50%	100,00%
71,30	97,00%	100,00%	97,75%	99,83%
95,00	95,00%	99,00%	67,85%	100,00%
167,30	71,05%	95,33%	75,50%	90,22%
191,30	69,70%	87,71%	78,62%	82,33%
215,30	68,00%	79,16%	68,88%	85,33%
239,00	55,03%	55,42%	77,00%	84,16%
263,00	59,37%	49,50%	77,86%	68,25%

	Exp 3	Exp 3	Exp 3	Exp 3
Time (h)	<b>Cirm 9</b>	<b>Cirm 121</b>	<b>Cirm 127</b>	<b>Cirm 138</b>
0,00	98,0%	100,0%	100,0%	100,0%
24,00	100,0%	92,0%	100,0%	89,5%
30,00	98,76%	99,23%	88,07%	99,08%
47,00	97,27%	100,00%	100,00%	99,61%
53,00	100,00%	99,25%	98,58%	99,90%
71,00	100,00%	100,00%	100,00%	99,00%
95,00	93,27%	78,50%	95,80%	87,16%
168,00	95,20%	72,60%	76,33%	88,16%
192,00	85,33%	87,00%	83,66%	78,83%
216,00	84,57%	81,58%	77,58%	78,60%
240,00	71,50%	84,33%	51,85%	49,00%
263,30	61,60%	77,00%	57,83%	61,50%

Time (h)	Exp 4	Exp 4	Exp 4	Exp 4
	<b>Cirm 1</b>	<b>Cirm 119</b>	<b>Cirm 129</b>	<b>Cirm 139</b>
0,00	97,7%	97,0%	98,5%	96,0%
23,00	100,0%	98,5%	100,0%	100,0%
29,00	100,00%	98,80%	95,00%	99,00%
47,00	100,00%	99,00%	94,71%	98,50%
53,00	100,00%	99,66%	98,83%	100,00%
71,00	98,42%	98,00%	100,00%	100,00%
94,30	98,33%	87,50%	95,60%	97,83%
166,00	97,83%	61,50%	69,58%	66,33%
189,30	69,00%	59,50%	63,00%	90,66%
212,00	67,00%	71,75%	74,70%	90,66%
235,30	39,88%	68,33%	78,87%	91,08%
259,00	62,00%	68,16%	65,85%	77,16%

**SEÇÃO II: Utilização da genômica funcional para a compreensão do fenômeno de sobrevivência à longo prazo em *Propionibacterium freudenreichii***

**SECTION II : L'utilisation de la génomique fonctionnelle pour comprendre le phénomène de la survie à long terme chez *Propionibacterium freudenreichii***



## 1. Introdução

Esta segunda parte dos resultados é também apresentada sob forma de artigo científico. Este artigo segue os resultados obtidos no estudo da caracterização fenotípica apresentada na seção I de resultados deste manuscrito de tese.

O objetivo deste estudo foi conhecer o metabolismo e o transcriptoma de *P. freudenreichii* antes e depois da entrada em fase estacionária para compreender sua adaptação fisiológica à sobrevivência a longo prazo. Durante a cinética de crescimento e a sobrevivência durante 11 dias (Secção I - Resultados) da linhagem CIRM-BIA138, o sobrenadante da cultura foi amostrado para dosar as quantidades residuais de lactato, lactose, glicose, frutose, manose do sobrenadante de cultura bem como de produtos da fermentação: piruvato, propionato, acetato e succinato. Em paralelo, o RNAm destas células foram extraídos e sequenciados através da técnica de RNA-Seq à 1, 3 e 9 dias após a inoculação.

A análise dos dados de sequenciamento por mapeamento foi possível apenas nos pontos de 1 dia e 3 dias após a inoculação, correspondentes à fase exponencial e a entrada na fase estacionária, respectivamente. Isto ocorreu porque por motivos de qualidade e quantidade de ARN, não foi possível obter resultados das amostras de 9 dias. A análise dos genes diferencialmente expressos por EdgeR mostrou que 912 genes foram expressos diferencialmente entre a fase exponencial e o início da fase estacionária do crescimento bacteriano, entre eles, 458 genes são induzidos e reprimidos 454.

Para a obtenção do perfil metabolômico desta linhagem, os dados de expressão gênica global de *P. freudenreichii* CIRM-BIA138 foram combinados aos dados obtidos pela dosagem de metabólitos durante todo este período de 11 dias.

Os resultados bioquímicos demonstram que *P. freudenreichii* CIRM-BIA138 utiliza toda a glicose, manose e frutose no início da fase exponencial de crescimento até a exaustão destes compostos. Esta consumação é seguida de uma libertação e recaptação de piruvato, metabolito intermediário destes açúcares. Durante esta mesma fase de crescimento, observa-se um consumo de lactato até a exaustão à entrada na fase estacionária e uma produção de propionato e de acetato em quantidade equimolar. Os dados de RNA-Seq demonstraram uma redução da expressão nos processos de transcrição, tradução e secreção de proteínas à 3 dias após a inoculação (entrada em fase estacionária) em comparação com a fase exponencial. Isto pode acontecer devido a uma carência de lactato no início da fase estacionária. Além disso, a análise de vias

metabólicas diferencialmente expressas revelou que a glicólise, a fosforilação oxidativa, o ciclo de Wood-Werkman, a síntese de peptidoglicano e de ácidos graxos foram reprimidos à 3 dias de cultura. Em contraste, à 3 dias após a inoculação, o gene de síntese da cápsula de polissacárideo é induzida, assim como as vias de degradação de inositol, arabinose, asparagina, glicina, alanina, polifosfatos e, provavelmente, a via de degradação de putrescina ao GABA através do glutamato.

Assim, integrando neste estudo os dados transcriptômicos e metabolômicos, foi possível encontrar pistas sobre os mecanismos envolvidos na adaptação à carência nutricional que permitir a sobrevivência a longo prazo em *P. freudenreichii* CIRM-BIA138. As vias reprimidas indicam, provavelmente, uma redução da atividade metabólica desta bactéria para a sua entrada em dormência. Estes resultados abrem caminhos para outros estudos mais direcionados para compreender melhor os mecanismos envolvidos na estratégia de adaptação para a sobrevivência a longo prazo em *P. freudenreichii*.

## 1. Introduction

Cette deuxième partie des résultats est aussi présentée sous forme d'article scientifique. Cet article fait suite aux résultats obtenus dans l'étude de la caractérisation phénotypique présentée dans la section I de résultats de ce manuscrit de thèse.

L'objectif de cette étude était d'appréhender le métabolisme et le transcriptome de *P. freudenreichii* avant et après à entrée en phase stationnaire pour comprendre son adaptation physiologique à la survie à long terme. Au cours de la cinétique de croissance et de survie pendant 11 jours (section I - Résultats) de la souche CIRM-BIA138, le surnageant de culture a été échantillonné pour doser les quantités résiduelles de lactate, lactose, glucose, fructose, mannose du milieu de culture ainsi que les produits de la fermentation : pyruvate, propionate, acétate et succinate. En parallèle, l'ARNm des cellules a été extrait et séquencé (RNA-Seq) à 1, 3 et 9 jours après inoculation.

L'analyse des données de séquençage par mapping n'a été possible, pour des raisons de qualité et de quantité d'ARN que sur les points 1j et 3j après inoculation correspondant à la phase exponentielle et à l'entrée en phase stationnaire respectivement. L'analyse des gènes différentiellement exprimés par EdgeR a montré que 912 gènes étaient différentiellement exprimés entre la phase exponentielle et le début de la phase stationnaire de la croissance bactérienne, parmi eux, 458 gènes sont induits et 454 réprimés.

Pour obtenir le profil métabolomique de la souche, les données de l'expression génique globale de *P. freudenreichii* CIRM-BIA138 ont été combinées aux données obtenues par dosage de métabolites au cours de toute cette période de 11 jours.

Les résultats biochimiques montrent que *P. freudenreichii* CIRM-BIA138 utilise tout le glucose, le mannose et le fructose au tout début de la phase exponentielle de croissance jusqu'à épuisement de ces composés. Cette consommation est suivie d'un relargage puis d'une reconsommation de pyruvate, métabolite intermédiaire de ces sucres. Pendant cette même phase de croissance, on observe une consommation du lactate jusqu'à épuisement à l'entrée en phase stationnaire et une production de propionate et d'acétate en quantité équimolaire. Les données biochimiques de la phase exponentielle sont en accord avec celles décrites dans la bibliographie. Pendant la phase stationnaire et la survie à long-terme le pyruvate présent dans le milieu est progressivement consommé jusqu'à épuisement à 11 jours après inoculation. Les données RNA-Seq ont montré une réduction d'expression dans les processus de transcription, traduction et sécrétion de protéines à 3 jours après inoculation (entrée en

phase stationnaire) en comparaison avec la phase exponentielle. Cela peut se produire en raison de la carence en lactate au début de la phase stationnaire. De plus, l'analyse des voies métaboliques différentiellement exprimées a montré que la glycolyse, la phosphorylation oxydative, le cycle de Wood-Werkman, la synthèse de peptidoglycane et d'acides gras étaient réprimés à 3 jours de culture. En revanche, à 3 jours après inoculation, le gène de synthèse de capsule polysaccharidique est induit ainsi que les voies de dégradation de l'inositol, de l'arabinose, de l'asparagine, de la glycine, de l'alanine, des polyphosphates et probablement la voie de dégradation de la putrescine en GABA via le glutamate.

Ainsi, en intégrant dans cette étude les données transcriptomiques et métabolomiques, il a été possible de trouver des pistes concernant les mécanismes impliqués dans l'adaptation à la carence nutritionnelle qui permettraient la survie à long terme chez *P. freudenreichii* CIRM-BIA138. Les voies réprimées indiquent probablement une réduction de l'activité métabolique de cette bactérie en vue de son entrée en dormance. Ces résultats ouvrent la voie à d'autres études plus ciblées pour mieux cerner les mécanismes impliqués dans cette stratégie d'adaptation à la survie à long terme chez *P. freudenreichii*.

## **2. Artigo 2: Adaptation of *Propionibacterium freudenreichii* to long-term survival in case of gradual nutritional shortage**

**Flavia Figueira Aburjaile<sup>1,2,3</sup>, Marine Rohmer<sup>4</sup>, Hugues Parrinello<sup>4</sup>, Marie-Bernadette Maillard<sup>2,3</sup>, Eric Beaucher<sup>2,3</sup>, Gwénaële Henry<sup>2,3</sup>, Aurélie Nicolas<sup>2,3</sup>, Marie-Noëlle Madec<sup>2,3</sup>, Anne Thierry<sup>2,3</sup>, Sandrine Parayre<sup>2,3</sup>, Stéphanie-Marie Deutsch<sup>2,3</sup>, Muriel Cocaïgn-Bousquet<sup>5,6,7</sup>, Anderson Miyoshi<sup>1</sup>, Vasco Azevedo<sup>1</sup>, Yves Le Loir<sup>2,3†</sup>, Hélène Falentin<sup>2,3†</sup>**

<sup>1</sup> Department of General Biology, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

<sup>2</sup> INRA, UMR 1253, Science et Technologie du Lait et de l'Œuf, 35000, Rennes, France

<sup>3</sup> AGROCAMPUS OUEST, UMR1253, UMR Science et Technologie du Lait et de l'Œuf, 35000, Rennes, France

<sup>4</sup> UMS BioCampus- MGX Montpellier GenomiX, Institut de Génomique Fonctionnelle, 141, rue de la Cardonnille, 34094, Montpellier Cedex 05, France

<sup>5</sup> Université de Toulouse, INSA, UPS, INP, LISBP, Toulouse, France

<sup>6</sup> INRA, UMR792 Ingénierie des Systèmes Biologiques et des Procédés, Toulouse, France

<sup>7</sup> CNRS, UMR5504, Toulouse, France

Corresponding author: Hélène Falentin, INRA, UMR 1253, Science et Technologie du Lait et de l'Œuf, Rennes, France, e-mail: [helene.falentin@rennes.inra.fr](mailto:helene.falentin@rennes.inra.fr).

†: YLL and HF share credit for senior co-authorship

## **Abstract**

### **Background**

*Propionibacterium freudenreichii* is an Actinobacterium widely used in the dairy industry as a ripening culture for Swiss-type cheeses, for vitamin B12 production and some strains display probiotic properties. It is reportedly a hardy bacterium, able to survive the cheese-making process and digestive stresses.

### **Results**

During this study, *P. freudenreichii* CIRM-BIA 138, which has a generation time of 5H in Yeast Extract Lactate medium at 30°C under microaerophilic conditions, was incubated for 11 days (9 days after entry into stationary phase) in a culture medium, without any adjunct during the incubation. The carbon and free amino acids sources available in the medium, and the organic acids produced by the strain, were monitored throughout growth and survival. Although lactate (the preferred carbon source for *P. freudenreichii*) was exhausted three days after inoculation, the strain sustained a high population level of  $9.3 \log_{10}$  CFU/mL. Its physiological adaptation was investigated by RNA-seq analysis and revealed a complete disruption of metabolism at the entry into stationary phase as compared to exponential phase.

### **Conclusions**

*P. freudenreichii* adapts its metabolism during entry into stationary phase by down-regulating oxidative phosphorylation, glycolysis, and the Wood-Werkman cycle by exploiting new nitrogen (glutamate, glycine, alanine) sources, by down-regulating the transcription, translation and secretion of protein. Utilization of polyphosphates and rearrangement at chromosome level were suggested.

**Keywords:** *Propionibacterium freudenreichii*, long-term survival, stationary phase, RNA-seq, adaptation.

## Background

*Propionibacterium freudenreichii* is an Actinobacterium widely used in the dairy industry and responsible for aroma development and opening in Swiss-type cheeses. Some strains are also used as probiotics because they produce bifidogenic compounds [1], they are resistant to digestive stress [2] and they may be endowed with anti-inflammatory capabilities and could be used to prevent inflammatory bowel diseases [3] [4]. *P. freudenreichii* core metabolism leads to propionic acid as the main end-product [5] [6].

The genetic bases for hardiness have already been well-documented. *P. freudenreichii* can produce ATP and NAD(P)H via fermentation utilizing an unique metabolic pathway (the Wood-Werkman cycle), leading to propionic acid. Under anaerobiosis, *P. freudenreichii* can also use electron acceptors other than O<sub>2</sub> (such as humic acid or nitrate) to produce ATP during anaerobic respiration [7] [8]. Several genes involved in polyphosphate and pyrophosphate utilization are found in the genome of CIRM-BIA1, the type strain for *P. freudenreichii*, suggesting a storage of energy in these forms [9]. However, these genes have never yet been found to be induced under any growth conditions.

*P. freudenreichii* can stay alive and metabolically active for long periods even under stressful conditions like gastro-intestinal tract environment or in cold condition. When placed at stationary phase in the intestinal tract (colon) of a piglet, CIRM-BIA1 was able to ferment some atypical carbon sources present in the gut (gluconate and propanediol), enabling its survival under these harsh conditions. Both of these catabolic pathways were reconstructed in silico from the genomic data and were found to be fully expressed in the colon of piglets [10]. Another transcriptomic study on the adaptation of

*P. freudenreichii* strains to the cold in the absence of carbon shortage showed that strains slowed down their cellular metabolism, displayed cold stress responses, and rerouted their carbon metabolism toward trehalose and glycogen synthesis [11]. In addition, proteomic studies on different *P. freudenreichii* strains have highlighted the role of chaperones during exposure to acids, bile salts or NaCl [12] [13] [14]. The genomic analysis of *P. freudenreichii* strains thus revealed the duplication of several chaperone genes.

In a previous study [15], we carried out the phenotypic characterization of *P. freudenreichii* over an 11-day period without the addition of nutrients, and revealed different phases of growth, membrane permeabilization and entry into dormancy and a viable but non-cultivable (VBNC) state to ensure Long-Term Survival (LTS). How *P. freudenreichii* prepares the LTS phase remains unknown and no data are currently available on the metabolism of *P. freudenreichii* in stationary phase in the event of a gradual nutritional shortage.

The objective of the present study was therefore to identify the pathways used by *P. freudenreichii* to cope with starvation and enter to the LTS phase. The strategy adopted here combined the quantification of sugars, acids and free amino acids in the supernatant of an 11-day culture without the addition of nutrients, and an RNA-seq analysis of bacterial cells sampled at three different time points during *P. freudenreichii* culture. This work was designed to compare both biochemical quantifications (acids, amino-acids and sugars) and transcriptomic data between the exponential and stationary phases in order to gain an in-depth view of *P. freudenreichii* adaptation to LTS.



## Results and Discussion

The objective of the study was to identify the metabolic pathways used by *P. freudenreichii* to cope with starvation and 'prepare' entry into the LTS phase.

### Enumeration, DO, pH

The results of bacterial enumerations and pH measurements (Figure 1A) confirmed that CIRM-BIA 138 survived well (around 8 log<sub>10</sub> CFU/mL) during 11 days, even under conditions of nutritional shortage, as it had been observed during a previous study [15]. When inoculated at 7 log<sub>10</sub> CFU/mL, the *P. freudenreichii* culture entered the exponential growth phase to reach a maximum population of 9.3 log<sub>10</sub> CFU/mL 3d post-inoculation, corresponding to entry into the stationary phase. The pH was very slightly affected by growth since the most acidic point was 6.7. At entry into stationary phase, a slight decline in the population was observed until it stabilized at around 8.5 log<sub>10</sub> CFU/mL as from 9d post-inoculation.

### Carbon source utilization, organic acid utilization and production

CIRM-BIA138 can degrade glucose, glycerol, mannose, galactose, inositol, erythritol, adonitol, esculine, lactose, xylitol, gluconate, D-fructose and L-arabinose [5]. In this study, CIRM-BIA138 was cultivated in Yeast Extract Lactate, a growth medium containing 120 mM lactate. Small quantities of several other carbon sources were also identified: 0.42 mM lactose, 0.15 mM glucose, 0.07 mM fructose, 0.03 mM mannose. Inositol and arabinose were not detected and probably below the detection limit of 0.5 mM. *P. freudenreichii* CIRM-BIA138 used lactose, glucose, mannose and fructose at the beginning of growth until they were completely exhausted at 24h post-inoculation (Figure 1B). These degradation capacities are in line with those previously described for

this strain [5]. The degradation of these sugars was concomitant with the release of pyruvate into the medium, this being one of their intermediary metabolites. From 24h post inoculation onwards, pyruvate was consumed by the strain. Levels of other compounds such as propionate, acetate and succinate rose gradually in the medium over time during growth, to reach a plateau at 3d post-inoculation (Figure1C). The stoichiometry of the conversion was 3 mol of lactate into 1 mol of acetate and 2 mol of propionate, and is in line with those described previously [16]. Lactate levels decreased as early as the onset of growth because of its consumption by *P. freudenreichii*. Lactate was exhausted 3d post-inoculation and pyruvate 11d post-inoculation. Among all the carbon sources quantified during this experiment, pyruvate was the only carbon source detectable in the medium when *P. freudenreichii* entered the stationary phase.

Overview of differentially expressed genes between exponential and stationary phases

After filtration to eliminate reads of poor quality, sequencing generated 100,450,770 reads (FastQC results), which included 13,502,207 non-aligned reads and 37,018,311 reads aligned on coding sequences. The remaining reads corresponded to those mapping on two different sequences (1,147,269 ambiguous reads according an htseq-count with a union parameter), on tRNA, rRNA, or non coding RNA (small RNA, non-sense RNA, 5' and 3' UTR). Only reads mapping on coding sequences were analysed further. Overall, 912 genes were found to be differentially expressed (adjusted pvalue<0.05) between 1d (exponential phase) and 3d post-inoculation (start of stationary phase) (458 genes induced (Additional file 1: Table S1), 454 genes repressed (Additional file 2: Table S2)). Differentially expressed genes (adjusted pvalues<0.05, pvalues available in Table S1, S2 and S3) represented 40% of the protein-coding genes in the CIRM-BIA138 genome. The

genes most affected by entry into stationary phase was those implicated in carbohydrate metabolism with 26 genes of this functional category being induced and 41 repressed. Translation was the second most affected category, with two genes induced and 56 repressed. Transport/binding was the third most affected category, with 28 genes induced and 19 repressed. Transcription was the fourth most affected category, with 29 genes induced and 17 repressed. Membrane bioenergetics was the fifth most affected category, with nine genes induced and 36 repressed. Amino acid metabolism category was also affected, with 11 genes induced and 32 repressed. Lastly, the metabolism of coenzyme and prosthetic groups was also affected, with 11 genes induced and 29 repressed (Figure 2). Most induction and repression observed from RNAseq analysis were confirmed by RT-qPCR analysis ( $p$ value $<0.05$ , Table S3). To validate RNAseq results, we performed RT-qPCR and we have included the 9d post inoculation sample. In most cases, the tendency (induction versus repression) seen at 3d post inoculation was confirmed at 9d. A massive number of differentially expressed genes (63 repressed and 177 induced) were annotated as having an 'unknown function'. 39% of induced genes at stationary phase were of unknown function. The genome of CIRM-BIA 138 contains 27 % of genes encoding protein of unknown function. So, results suggested that genes expressed at the entry into stationary phase are enriched in genes without function, thus reflecting our lack of knowledge on bacterial stationary phase metabolism. All induction and repression discussed below are statistically significant at adjusted  $p$ value $<0.05$  for RNAseq and at  $p$ value $<0.05$  for RT-qPCR.

#### Slowdown of the metabolism leading to dormancy

RNA extraction and quantification throughout growth and survival showed that RNA levels gradually declined. The total/depleted RNA quantities corresponding to 2 mL of

culture were 60µg/290 ng at 1d post-inoculation (exponential growth phase) for 15 log<sub>10</sub> CFU, 22 µg/225 ng at 3d post-inoculation (entry into stationary phase) 19 log<sub>10</sub> CFU and 15 µg/90 ng at 11d post-inoculation 18 log<sub>10</sub> CFU. The translation machinery was also reduced at 3d post-inoculation. Indeed, expression of the translation initiation factor IF-3, and the elongation factors Tu and G encoded by genes *infC*, *tuf*, *fusA*, were 2, 1.8 and 3-fold repressed, respectively. The expression of ribosomal proteins was repressed from 1.5 to 16 fold (see Table S2, for repression fold and pvalue of each gene). Taken together, these repressions reflected a slowdown of the cellular machinery in line with the process of dormancy recently described [15]. This slowdown can be explained by the scarcity of carbon sources when *P. freudenreichii* enters the stationary phase at 3d post-inoculation. The main glycolytic pathways producing reducing equivalents were repressed during this phase. In YEL medium, glycolysis, the pentose phosphate pathway and the Wood-Werkman cycle (fermentation) are three ways for *P. freudenreichii* to produce NADH and NADPH reducing equivalents. ATP and precursor metabolites are required for the biosynthesis of essential compounds (amino acids, purine, pyrimidine, glycerol 3 phosphate, fatty acids, N-acetyl glucosamine, vitamins). Here, under the conditions we applied, glucose was exhausted at 3d post-inoculation, and the genes involved in glycolysis were repressed. Some of them were particularly strongly down-regulated, such as *sglT*, the glucose transporter with a fold change of 4, and *pfkA*, *pfp*, *fba1* and *eno2* with fold changes of 2.2, 3.1, 1.9 and 3.4, respectively. The repression of *sglT* was confirmed by RT-qPCR results (pvalue<0.05) with a repression fold of 6 at 3d and 88 at 9d post-inoculation (Table S3).

The Wood-Werkman cycle is specific to certain propionic acid-producing bacteria. It plays a central role in propionic acid fermentation, the principal carbon metabolic pathway in dairy propionibacteria. This pathway provides major sources of reducing

equivalents: NADPH required for biosynthesis reactions and NADH necessary for oxidative phosphorylation. Through this pathway, pyruvate is converted into propionate. Pyruvate is initially converted into succinate by the successive steps of the tricarboxylic cycle (TCA). At 3d post-inoculation, gene-set enrichment analysis identified the “Citrate Cycle” as one of the most down-regulated pathways (ko00020, fold change of 2.1, pval=2.7e-04). All the corresponding genes (mdh, fumC and sdh) were repressed, with fold changes of 1.9, 2.1 and 8.6, respectively, in stationary phase compared to exponential phase (see Table S2). Succinate is then converted into succinyl-CoA, methyl malonyl CoA, propanoyl CoA and propionate by specific enzymes. Transcripts corresponding to the 12S, 5S and 1.3S subunits of the well-studied methylmalonyl-CoA carboxyltransferase were repressed with fold changes of 2.3, 1.8 and 2.9, respectively. RT-qPCR results confirmed the repression of mmda, encoding the 12S subunit with a fold change of 2. Down-regulation of the Wood-Werkman cycle is probably involved in maintaining the redox balance. Although glycolysis and the Wood Werkman cycle were down-regulated, all genes in the pentose phosphate pathway (except the rpiB3 gene, induced with a fold change of 2.9) were stably expressed between 1d and 3 d post-inoculation.

#### Changes in oxidative phosphorylation

Several transcripts (nuoA,B,C,D,E,F,I,J,K,L,M,N genes) encoding the different chains of NADH-quinone oxidoreductase (responsible for the release of electrons and H<sup>+</sup> contained in the NADH molecule) involved in aerobic respiration were down-regulated at 3d post-inoculation, with a fold change of between 3.8 and 6.3. The PFCIRM138\_05930 gene encoding an iron-sulphur protein was also down-regulated, with a fold change of 6 in RNAseq analysis and 4.5 at 3d and 5.5 at 9d post- inoculation according to RT-qPCR results. Likewise, sdhA, A3, B, B3, C1, and C2, encoding the

different subunits of succinate dehydrogenase, were down-regulated with fold changes of between 4.4 and 8.8 (for details, see Supplementary Table S1 and S2). Repression of *sdhC1* was confirmed by RT-qPCR results with a repression with a fold of 9 (Table S3). By contrast, both the *dmsC* and *dmsB* genes, encoding the anaerobic dimethyl sulphoxide reductase chains C and B, were induced with fold changes of 13.7 and 4.7, respectively. The latter protein handles the final transfer of electrons to various sulphoxide and N-oxide compounds. Anaerobiosis-inducible dimethyl sulphoxide reductases play a key role in bacterial adaptation to anaerobic conditions in bacteria and serve as terminal reductases using DMSO as a terminal electron acceptor [17]. Nitrate, sulphur or ferrous ions can act as electron acceptors. Accordingly, the gene encoding the permease protein of the nitrate ABC transporter *ssuC* was induced with a fold change of 4.4, the gene *citT2* encoding a sodium:sulphate symporter was induced with a fold change of 4.3, and the *feuC* and *feuS* genes encoding the ferrous ABC transporter were induced with fold-changes of 2.5 and 2.8. Some subunits of nitrate reductase, namely those encoded by *narH* and *narJ*, were induced by more than 22-fold. RT-qPCR results confirmed the induction of *narH* with a lesser fold of 3.5. During growth, the headspace of the tube contained air and was therefore a source of O<sub>2</sub>. Although the culture was grown without agitation, the medium contained traces of O<sub>2</sub>, which were probably used as a terminal electron acceptor during the exponential growth phase. At 3d post-inoculation, the CO<sub>2</sub> released by the bacteria likely saturated the headspace of the tube and the *P. freudenreichii* were in anoxic conditions. This might modify the redox balance due to the lower availability of reducing equivalents, and thus explain the down-regulation of oxidative phosphorylation. At the pathway level, gene-set enrichment analysis (see Materials and Methods) showed that oxidative phosphorylation was indeed the most significantly down-regulated (ko00190, fold change of 4.4 and

pval=1.2e-12) of all of the pathways down-regulated during the study. A similar down-regulation had been observed when *P. freudenreichii* was placed in the colon of piglets under anoxic conditions [10]. To assess whether the oxydative repression is due to traces of O<sub>2</sub> in the headspace or due to the metabolism at the entry into stationary phase, cultures growing in bioreactors under strictly controlled conditions can be considered.

The changes thus observed suggest profound metabolic reprogramming in response to starvation. Similarly, during our study, specific catabolic pathways were induced (see below).

### Cell wall

The cell envelope of Gram positive bacteria such as *P. freudenreichii* comprises the inner cell (cytoplasmic) membrane and the cell wall of the bacterium composed of peptidoglycan and various associated compounds (proteins, polysaccharides, teichoic acids) that differ from one species to another. In *P. freudenreichii*, some strains are known to possess a surface exopolysaccharidic (EPS) layer composed of (1 → 3,1 → 2)-D-glucan [18].

Peptidoglycan (PG) is essential to maintaining cell shape and also providing mechanical protection against osmotic pressure. It is also involved in cell division process. PG is a three-dimensional network made up of N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid(MurNAc), The carboxyl groups of MurNAc are substituted by a short peptides, interconnecting the chains together. The biosynthesis of PG includes cytoplasmic steps mediated by MurA to F, and thereafter the transfer through the membrane via the transferases MarY and MurG, and finally the binding of new material to cell wall. At entry into stationary phase,

the PG biosynthesis pathway is significantly down-regulated. The genes encoding MurA and MurB that are involved in the formation of UDP-N-acetylmuramate were down-regulated, with a fold changes of 1.8 for both. The murC, D, E, and F genes encoding cytoplasmic enzymes responsible for the sequential adjunction of amino-acids to UDP-N-acetylmuramate, leading to the formation of UDP-MurNac-pentapeptide, were also down-regulated (with fold changes of 2.3, 1.6, 1.2, and 2, respectively) as well as the gene *ddlA* *reponsible for the formation of the D-alanyl-D-alanine dipeptide is down regulated (fold changes of 2,9)*. Two proteins MraY and MurG ensure the transfer of the phospho--MurNac-pentapeptide moiety of the UDP-MurNac-pentapeptide to the membrane acceptor and the addition of GlcNac, leading to the formation of GlcNac-MurNac-pentapeptide on the lipid carrier. The *mraY* gene encoding UDP- MurNac-pentapeptide phosphotransferase was repressed with a fold change of 3.4, whereas expression of the *murG* gene remained stable. Finally, the enzymes responsible of the late steps of the biosynthesis of the PG are also down regulated: *fstI* (fold change of 1,9), *mrc/ponA* (FC of 1,5) Taken together, our results showed a down-regulation of peptidoglycan synthesis, which agreed with the growth arrest and entry into dormancy observed for CIRM-BIA138.

The CIRM-BIA138 strain produces a surface -D-glucan polysaccharide [18]. A single *gtfF* gene is responsible for synthesis of this polysaccharide. The *gtfF* gene is strongly induced in stationary growth phase compared to the exponential phase, with a fold change of 6.1, which might lead to an overproduction of surface -D-glucan polysaccharide. The induction was confirmed with a fold change of 3.3 at 3d but not at 9d post-inoculation by RT-qPCR results. Such EPS production might protect the bacteria against the unfavourable conditions encountered in stationary phase and



therefore prepare them for long-term survival, as has previously been shown in other food species such as *Oenococcus oeni* [19].

The cell wall of *P. freudenreichii* is also coated with proteins that are anchored via an SLH domain. SlpE, SlpF and SlpG are three of these proteins. Interestingly, during this study, the *slpE* gene was found to be strongly up-regulated (fold change of 12.9); whereas *slpG* and *slpF* were down-regulated (fold change of 2.8 and 2.2 respectively). According to RT-qPCR results, the induction of *slpE* is transient since it was confirmed at 3d with a fold change of 5.5 but not at 9d post-inoculation. As of yet, the physiological role of these proteins has not been elucidated, but these results suggest differing roles in long-term survival for SlpE and the other two proteins.

The bacterial membrane mostly comprises a protein-embedded phospholipid bilayer. In CIRM-BIA138, the fatty acid biosynthesis pathway was strongly repressed in stationary phase compare to exponential phase. The *fabF* and *fabH* genes encoding the enzyme charging the acetyl residue from acetyl-CoA to acyl-carrier protein (Acp) were repressed, with respective fold changes of 6.3 and 6.9. RT-qPCR results confirmed repression of *fabF* with a fold-change of 8.2 at 3d and 6.6 at 9d post-inoculation. The *acp* gene was repressed, with a fold change of 5.9. In the same way, the *fabD* genes encoding the enzyme charging the malonyl residue from malonyl-CoA to Acp was repressed, with a 3.7-fold change. Genes such as *fabG* or *inhA*, which are responsible for the further elongation of fatty acids, were repressed by 1.9 and 1.7, respectively. Bacteria can produce fatty acids anaerobically. Since most fatty acids in bacterial cells are used for membrane phospholipid synthesis, growth arrest at the entry into stationary phase at 3d

compared to 1d post-inoculation (see CFU count, Figure 1) might limit the need for fatty acids, and lead to the down-regulation we observed here.

### Diversification of nutrients

At 3d post-inoculation, glucose, lactose, and lactate were exhausted. *P. freudenreichii* therefore needed to recruit other catabolic pathways to produce NADH and NADPH reducing equivalents, ATP and precursor metabolites required for the survival. Inositol and arabinose pathways were induced. Gene-set enrichment analysis revealed that 'Pentose and glucuronate interconversions' were the most markedly induced KEGG pathways (ko00040; fold change of 1.7, pval 0.003). In this large pathway, the degradation of arabinose appeared to be the only one to be induced.

### Arabinose

Most *P. freudenreichii* strains can degrade L-arabinose, and the degradation pathway has previously been described [5]. L-arabinose enters the cell via a xylose/ribose/arabinose/galactoside ABC transporter encoded by rbsBA that is not differentially expressed. L-arabinose is sequentially converted to L-ribulose, L- ribulose 5-phosphate, and D-xylulose 5-phosphate by the action of the L-arabinose isomerase encoded by araA, L-ribulokinase encoded by araB and ribulose-5- phosphate 4-epimerase encoded by araD and araD1, which in our study were induced 2.2, 2.2, 1.1 and 2.9 fold, respectively. However, biochemical quantification failed to detect arabinose which is probably present at a concentration lower than the detection limit of 0.5 mM (like glucose, lactose, fructose, mannose) in the medium.

### Inositol

Inositol is a six-fold cyclohexane alcohol. Inositol is found in many foods (particularly in fruits) and is probably present in the yeast extract contained in YEL culture medium. Inositol is transported into the cell by a transporter encoded by *iolT1*, *iolT2* and *iolT3*, which were found to be induced with respective fold changes of 11, 3.2 and 1.7, at the entry into stationary phase compared to exponential phase (Table S2). The induction of *iolT1* was induced 2.4 at 3d and 58.3 at 9d according to RT-qPCR results. In the cell, inositol is transformed into 2-keto-inositol by inositol dehydrogenase that is encoded by the *iol*, *iolG2* genes, induced with fold changes of 5.5 and 2.1, respectively. 2-keto-inositol is then transformed into 2,3-di-keto-4-deoxy inositol by 2-keto inositol dehydratase encoded by *iolE3*, which was found to be induced by a 2.2 fold change. 2,3-di-keto-4-deoxy inositol is then transformed into 2-deoxy-5-keto gluconic acid by the product of the *iolB* gene repressed with a fold change of 2.1. The *iolC* and *iolD* genes, whose products convert 2-deoxy-5-keto gluconic acid into malonic semialdehyde, were not differentially expressed. The induction of inositol degradation was not detected by GAGE, probably because the steps in this pathways are not fully described in the KEGG orthology map for *P. freudenreichii* species. The inositol degradation pathway had previously been annotated manually in all sequenced strains able to degrade inositol [20], and this revealed a high number of paralogous genes at different steps in the pathway (Figure 3), suggesting that the pathway is essential for this species. During our study, we were able for the first time to demonstrate the expression and induction of the inositol degradation pathway in *P. freudenreichii*. Unfortunately, inositol was undetected in the culture medium (data not shown).

### Use of asparagine

Biochemical analysis of the culture medium showed that asparagine was totally consumed between 1d and 3d post-inoculation (1.3 mM consumption, Figure 4). The KEGG and Metacyc pathways report that in bacteria, asparagine is degraded into aspartate that is further transformed into fumarate or succinate. Despite asparagine exhaustion in the CIRM-BIA138 supernatant, the transcripts of *asmA* and *asmB* enabling the transformation of asparagine into aspartate were both repressed at 3d post-inoculation compared to 1d post-inoculation, with fold changes of 2.1 and 1.5, respectively. In *P. freudenreichii*, the fate of asparagine between 1d and 3d post-inoculation remains unknown [21]. The use of asparagine as an energy supply has previously been suggested [18]. Asparagine is described as being co-metabolized along with aspartate and lactate in type strain *P. freudenreichii* CIRM-BIA1 [22]. However, the quantification of free aspartate did not support this idea in the *P. freudenreichii* CIRM-BIA138 strain. The concentration of free aspartate in the medium was not significantly different between 1d and 3d post-inoculation (t-test, P value >0.05). The majority (i.e. 70% of the 100 *P. freudenreichii* isolates previously tested) displayed very low levels of aspartate activity [22]. In CIRM-BIA138, the lack of aspartate consumption could be explained by the lack of the *dcuA* gene enabling the transport of aspartate into the cell.

### Putrescine

Putrescine is a deleterious biogenic amine that is often present in mature cheeses. In general, Swiss-type cheeses containing the *P. freudenreichii* species do not exhibit high levels of putrescine [24]. Whether this is due to the degradation of putrescine by *P. freudenreichii* remains to be demonstrated. The degradation pathway for putrescine has

not yet been described in *P. freudenreichii*. During the present study, the PFCIRM138\_100015 gene, putatively encoding a putrescine importer, was induced with a fold change of 3.7 at entry into stationary phase. In *E. coli*, putrescine degradation is supported by the *gadA* and *gadB* gene products and involves steps that are shared with those of glutamate. The glutamate decarboxylase encoded by *gadB* was 5-fold more induced at 3d compared to 1d post-inoculation. RT-qPCR results confirmed a strong and progressive induction of *gadB* with a fold change of 3 at 3d and of 114 at 9d post-inoculation. The glutamate concentration did not differ statistically at 3d and 1d post-inoculation (t-test, P value >0.05), suggesting that degraded glutamate arose from the prior degradation of putrescine. This induction was also consistent with that observed in *E. coli* for the *gadA* and *gadB* genes during entry into stationary phase. It has been shown that glutamate decarboxylase activity increases the survival of *Listeria monocytogenes*, because intracellular glutamate is stored and converted into  $\gamma$ -aminobutyric acid (GABA), and extracellular glutamate is coupled to an antiporter in the *gad* system [25]. The CIRM-BIA138 genome contains two genes encoding a GABA transferase: PFCIRM138\_10030 (*gabT*), which was repressed with a fold change of 6.6, and PFCIRM138\_04365, which was induced with a fold change of 10.6. However, neither GABA nor putrescine were detected in the medium despite a quantification limit of 25  $\mu$ M (data not shown). The induction of putrescine importer (PFCIRM138\_100015) and *gadB* at d3 and d9 is not sufficient to know whether these both genes were involved in the metabolism of stationary phase and whether glutamate metabolism could be linked to an hypothetical putrescine degradation by *P. freudenreichii*.

### Use of glycine

The dosage of free glycine in the medium revealed a 0.42 mM consumption between 1d and 3d post-inoculation (Figure 4). The *dadA2* gene encoding glycine oxidase, which catalyses glycine oxidative deamination was induced at 3d compared to 1d post-inoculation, with a fold change of 13 according to RNAseq analysis but the induction was not statistically significant according to RT-qPCR result. This gene is the only one revealing a discrepancy between RNAseq data and RT-qPCR analysis. However we can conclude to a maintain of *dadA2* expression at 3d and even at 9d post-inoculation. During osmotic stress induced by a high NaCl concentration in a medium containing glycine betaine (also named trimethylglycine), it was found that *Propionibacterium acidipropionici* accumulates glycine betaine in order to balance osmolarity and enable adaptation [26]. Other studies have also shown that in *P. freudenreichii*, glycine and proline exert protective effects against high osmolarity throughout bacterial growth [21]. Internalization of glycine could be a way to cope with osmotic conditions encountered at the entry of stationary phase.

### Use of alanine

Alanine was the free amino acid that was most consumed (1.17 mM between 1d and 3d post-inoculation). This was in line with induction of the *ald* gene encoding an alanine dehydrogenase (2.7 fold change), which deaminates L-alanine into pyruvate. The induction of *ald* throughout growth has already been observed in stationary phase induced by low temperature in the absence of nutrient shortage [6]. We showed here that such *ald* induction also occurred in a context of nutrient shortage.

### The role of proline

Biochemical analysis between 1d post-inoculation and 3d post-inoculation an almost two-fold increase in proline was measured in the supernatant (0.87 mM released) (Figure 4). As no gene for proline synthesis was induced at 3d post-inoculation, this suggests that proline was released due to the degradation of proteins and peptides in the medium. This was consistent with the action of a proline iminopeptidase (Pip) which cleaves amino-terminal proline residues from peptides and has previously been described in cheese [27].

### Chaperone and detox

Protein chaperones protect other proteins against aggregation and denaturation. In *P. freudenreichii*, chaperones are strongly expressed in the case of acidic or bile salt stresses [31] or under cold conditions [6]. At 3d post-inoculation, the three sequences hsp20 1, hsp20 2 and hsp20 3 encoding heat shock proteins were induced with fold changes of 6.7, 6.9 and 2.5, respectively. The copZ gene encoding a copper chaperone was induced with a fold change of 2.5. The cspB gene encoding a cold shock protein was induced with a 2.7 fold change. The number of chaperones induced seems low when compared to the large number of chaperone-encoding genes described in the literature for *P. freudenreichii*. The limited induction of chaperone genes might reflect the efficient adaptation of *P. freudenreichii* to nutritional shortage. Alternatively, and unlike osmotic or heat-shock stress, nutritional shortage may not induce protein misfolding or denaturation and thus not cause the expression of many chaperone genes.

## Competence

In several species, natural competence enables the bacteria to cope with stresses or nutritional shortage. There are three prevailing hypotheses regarding the benefits of DNA uptake and recombination in bacteria (for a review, see [32] [33]): (i) DNA is used as a nutrient source; (ii) DNA is used to improve the efficiency of natural selection (by acquiring new genetic information); and (iii) DNA is used as a template to repair damaged chromosomal DNA. In the case of *P. freudenreichii*, a gene encoding DNA transfer protein (PFCIRM138\_00630) was induced with a fold change of 14.7, and dprA encoding a DNA processing /uptake protein was induced with a fold change of 27.3 at 3d post inoculation. Two contiguous genes encoding mobilization protein (PFCIRM138\_07015, PFCIRM138\_07020) were induced with fold changes of 49.4 and 10.8 respectively, one gene (PFCIRM138\_07570) encoding a conjugative relaxase was induced with a 24-fold change and lastly the comEC gene involved in competence was induced with a fold change of 17.4 at 3 d post-inoculation. PFCIRM138\_07015 induction was confirmed by RT-qPCR result with a fold change of 23.6 at 3d and 26.7 at 9d post-inoculation. *P. freudenreichii* is known to have evolved more by recombination than by mutations [34] and possesses a high copy number of insertion sequences and integrase genes in its genome [9]. Experiments with DNA shuffling performed to improve the production yields of vitamin B12 also showed a high capacity for genome rearrangement in *P. freudenreichii* [35]. Here, several genes involved in natural competence were found to be induced under the conditions applied. However, neither natural competence nor conjugation has been described in this species to date. Cultures of CIRM-BIA138 in presence of high concentration of insertional plasmid (data not shown) failed to illustrate natural competence.



### Utilization of pyrophosphate

*P. freudenreichii* can accumulate inorganic polyphosphate (polyP) as an energy reserve whereas most bacteria utilize ATP [9]. The ability to use polyphosphate as an energy reserve has been shown to be specific to bacteria adapted to difficult environments. The CIRM-BIA138 genome possesses 20 genes encoding enzymes which use polyphosphate or pyrophosphate. At the start of the stationary phase during our study (at 3d post-inoculation), three phosphorylases were induced when compared to exponential growth phase at 1d post-inoculation: (i) *ppx5*, encoding an exopolyphosphatase with a 19.8 fold change, (ii) *ppa* encoding an inorganic pyrophosphatase with a 1.6 fold change, and (iii) PFCIRM138\_07685 encoding a NUDIX hydrolase with a 3 fold change. The *phoH* gene, encoding a phosphate-starvation inducible protein, was induced with a 2.5 fold change. In *Corynebacterium glutamicum*, another Actinobacterium that is phylogenetically close to *P. freudenreichii*, previous microarray results showed an induction of the *phoH* gene with a fold change ranging from 1.1 to 6.8 under Pi-limiting conditions compared to non-limiting conditions [28]. In many species, strains with a mutation in the gene involved in polyphosphate (*ppk*) synthesis are unable to survive during stationary phase [29]. Taken together, these results suggested a possible phosphate exhaustion of the medium, causing *P. freudenreichii* to use polyphosphates stored in its cytoplasm during stationary phase to cover its phosphate requirements. Since the phosphate starvation response is critical for the persistence of *Mycobacterium tuberculosis* (another Actinobacterium) in the lung [30], the role of phosphate starvation in the entry of *P. freudenreichii* into dormancy needs to be further explored, e.g. by quantifying extracellular phosphate and intracellular polyphosphate and by directed mutagenesis on key enzymes in the polyphosphate synthesis pathway.

## Conclusions

At entry into stationary phase, *P. freudenreichii* adapts its metabolism to nutritional shortage and slows down its metabolism. Genes involved in oxidative phosphorylation and fermentation (via the WoodWerkman cycle) are repressed, in line with the lack of lactate in the medium, enabling the entry into dormancy. By contrast, *P. freudenreichii* diversifies its source of nutrients and appears to utilize amino acids which differ from those used during the exponential phase according to RNAseq analysis, RT-qPCR and amino acids quantification. To meet its energy needs *P. freudenreichii* probably utilizes polyphosphate, because several phosphatases were found to be induced during this study. These results therefore provide an analysis of *P. freudenreichii* adaptation during entry into stationary phase by means of comprehensive gene expression analysis using RNA-Seq combined with targeted biochemical quantifications. They shed light on important molecular mechanisms that might be involved in the long-term survival of *P. freudenreichii*, and open avenues for further investigation on its survival strategies.

## Methods

### Bacterial strains

*P. freudenreichii* CIRM-BIA138 strain was used during this work. The strain was supplied by the International Centre for Microbial Resources-Bacteria of Food Interest (Centre International de Ressources Microbiennes-Bactéries d'Intérêt Alimentaire; INRA, Rennes).

### Conditions for bacterial growth

The strain was cultured in YEL medium (pH=7.0) [36] with no agitation at 30°C for 11 days. The strain was grown in a different assay tube for each manipulation, thus preventing any interference of oxygen in the environment and maintaining the strain under microaerophilic conditions. At T0, the medium was inoculated at  $10^7$  CFU/mL.

The growth kinetics were followed using Optical Density measurements with a Model DU 640 spectrophotometer (Beckman Coulter, Fullerton, California, USA) at 650 nm ( $OD_{650}$ ), and the CFU (CFU/mL) were counted using the micromethod described previously [37]. *P. freudenreichii* enumerations were carried out on YEL agar at 30°C under microaerophilic conditions, until the visualization of colonies (6 days). Two technical replicates and two biological replicates were performed for growth curves analysis, and three biological replicates were performed for RNA-seq and biochemical analysis. Aliquots of the culture were sampled for RNA extraction, pH measurements and biochemical analysis (see above) on the day of inoculation, and 24h (1 day), 72h (3 days) and 216h (9 days) post-inoculation.

### RNA extraction and quality control

1 ml of each culture was mixed with 2 volumes of RnaProtect (Qiagen, Hilden, Germany), left for 5 min. at room temperature and then centrifuged (8,000 g, 10 min at room temperature). The supernatant was removed and the pellet stored at -80°C until total RNA was extracted. The pellets were thawed on ice, suspended in 200 µL lysis buffer (50 mM Tris-HCl, 1 mM EDTA; pH 8.0) containing 20 mg/mL lysozyme (MP Biomedicals, Illkirch, France) and 50 U/mL mutanolysin (Sigma, Saint Quentin Fallavier, France), and incubated for 15 min at 24°C. The suspensions were then transferred to two millilitre tubes containing 50 mg zirconium beads (diameter: 0.1 mm; BioSpec Products, Bartlesville, OK) and 100 µL SDS (10%). The tubes were shaken twice for 90 s at 30 Hz with a bead beater (MM301; Retsch, Haan, Germany), being chilled on ice for 2 min. between the shaking steps. RNA extraction was then performed using an RNeasy minikit (Qiagen) and the Qiacube extraction robot (Qiagen), according to the manufacturer's instructions. RNA were suspended in 50 µL RNase-free water and treated with DNase (DNA-free; Ambion, Cambridgeshire, United Kingdom) according to the supplier's instructions, and then stored at -80°C until use. The RNA was then quantified and the contamination of RNA by proteins was assessed spectrophotometrically using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Inc., Rockland, DE). RNA quality was evaluated using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). All the RNA samples from spent medium had a RIN value higher than 7.5, indicative of good rRNA integrity. RNA was then depleted using RiboZero Magnetic kit for Gram positive bacteria (Epicentre, Madison, WI, USA) according to the manufacturer's instructions. Depleted RNA quality was evaluated using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) and nanoarrays for prokaryotes.

### Sequencing of mRNA

The cDNA libraries were prepared from the depleted RNA for each of the three repetitions of the exponential phase (1d post-inoculation), stationary phase (3d post-inoculation) and long-term survival (9d post-inoculation) and then prepared for sequencing using the Illumina TruSeq Stranded mRNA LT Sample Preparation kit. The nine oriented RNA-seq libraries were prepared using the TruSeq Stranded mRNA Sample Preparation kit, according to the manufacturer's instructions. Qualitative and quantitative library validations were performed on a DNA 1000 Agilent chip as well as using quantitative PCR with SYBR Green (Applied Biosystem 7500). The libraries were sequenced on an Illumina HiSeq 2000 as 50 bp reads using the Sequence By Synthesis technique at the Montpellier GenomiX facility. Image analyses and base calling were performed using the Illumina HiSeq Control Software and Real-Time Analysis component. Demultiplexing was performed using Illumina's sequencing analysis software (CASAVA 1.8.2). Data quality was assessed using FastQC from the Babraham Institute and the Illumina software SAV (Sequencing Analysis Viewer). The concentration in the samples corresponding to long-term survival (9d post inoculation) were too low for sequencing and revealed contamination with genomic DNA, so they were not further analysed.

### RNA-seq analysis

Sequences were mapped with Bowtie [38] and counted with htseq-count [39] using the ABIMS Roscoff platform. A list of differentially expressed genes was generated using an R software package: SARtools <https://github.com/PF2-pasteur-fr/SARtools> developed by Institut Pasteur Paris that embedded Deseq2 [40] and EdgeR [41] (adjusted Pvalues<0.05). A comparison of the results obtained with either Deseq2 or

EdgeR produced a more exhaustive list using EdgeR. Most of the differentially expressed genes from Deseq2 were present in the EdgeR results. For this reason, only the results obtained using EdgeR are discussed below.

The protein sequences of the strain were subjected to a search against the Kyoto Encyclopedia of Genes and Genomes Pathway database (KEGG) using GhostKOALA (<http://www.kegg.jp/ghostkoala/>) to retrieve KEGG orthology (KO) identifiers. 1143 out of 2304 genes were annotated automatically. Differentially expression data were associated with these KO identifiers and underwent gene set enrichment analysis on KEGG orthology pathways using the GAGE package [42] of R software to highlight pathways that were significantly differentially expressed (pval <0.1).

#### RT-qPCR validation

*cstA*, *eno1* and *sdaA* were used as housekeeping genes. Relative quantification relates the PCR signal of the target transcript at 3d and 9d post-inoculation to that of inoculation time as control. Before, RT-qPCR, a supplementary DNase treatment was applied to total RNA from 0, 3d and 9d post-inoculation samples. cDNA were obtained from a retrotranscription step using iScript cDNA synthesis kit (Bio-Rad, Marne la Coquette, France). cDNAs were amplified as follow: each PCR mixture included 5µl of 1/50 diluted cDNA, 3µl of 200 nM of primers, 8 µl of IQ Sybr Green supermix (Bio-Rad, Marne la Coquette, France). Amplifications were carried out on a CFX96 Real Time System (Bio-Rad) for 5 minutes at 95°C and 40 cycles of 2 steps consisting of 15 seconds at 95°C and 60 seconds at 60°C. The relative quantification of the mRNA levels of the target genes was determined using CFX Manager. The amount of target was normalized to *cstA*, *eno1* and *sdaA* genes because they were revealed as stably expressed according to the software:  $\Delta Cq = Cq(\text{target gene}) - Cq(\text{housekeeping})$

gene), where  $C_q$  represents the cycle number required to reach a defined threshold target abundance.

### Metabolome analysis

The samples used for this study had previously been frozen at  $-20^{\circ}\text{C}$ . These analyses were performed on the following samples: (i) YEL culture medium, (ii) exponential (1d post-inoculation), (iii) stationary (3d post-inoculation), (iv) late stationary (9d post-inoculation) and (v) long-term survival (11d post-inoculation), with three biological replicates of each sample. The YEL culture medium was used as a control.

#### ) Amino acid assays

Proteins and peptides in the samples were precipitated by adding solid sulphosalicylic acid to 5% (w/vol), holding at  $4^{\circ}\text{C}$  for 1 h, and centrifuged (5,000 g, 15 min). The supernatant was filtered through a  $0.45\ \mu\text{m}$  pore size filter and diluted with a 0.2 mol/L lithium citrate buffer (pH 2.2) before injection. Amino acid analyses were then carried out using cation exchange chromatography on a Biochrom 30 AA analyser (Biochrom Ltd, Cambridge, UK) according to the method described by Moore et al. (1958) with lithium citrate buffers as eluents and the ninhydrin post-column reaction system. Data were recovered using the control software Biosys and chromatographic data were processed with EZChrom Elite.

#### ) Quantification of lactate, pyruvate, acetate, citrate and propionate

Lactic, propionic, acetic, succinic and pyruvic acids were quantified in culture supernatants using an HPLC Aminex A-6 ion exchange column (Bio-Rad, Hercules,

CA) at 60°C with 0.005 M H<sub>2</sub>SO<sub>4</sub> at an eluent flow rate of 0.4 ml min<sup>-1</sup>. Acids were detected by UV (210 nm) and /or refractometry (RI2031 plus, Jasco).

- Quantification of arabinose and inositol

Concentrations were measured in culture supernatants by high performance liquid chromatography (Agilent Technologies 1200 Series, Waldbronn, Germany) using a HPX87H+ Biorad column and the following conditions: a temperature of 48 °C, eluent H<sub>2</sub>SO<sub>4</sub> (5 mM) at a flow rate of 0.5 mL/min, and dual detection (refractometer and UV).

#### J Quantification of glucose, mannose, fructose and lactose

The supernatant of culture medium was deproteinized using Vivaspin 10 kDa filters. Glucose, mannose, fructose, and lactose were separated on a CarboPacMA1 (4 x 250 mm) analytical column (preceded by a corresponding guard column 50 x 4 mm) with 16 mM NaOH as the eluent and a flow rate of 1 mL/min using an ICS-3000 chromatography/detector module (Dionex) and a range of 1 µC. A solution containing each of the four sugars (glucose, mannose, fructose and lactose) (Sigma-Aldrich) at 2, 5, 10, 20 and 40 mg/L (linearity range) was used as a standard. Detection and quantification were performed using amperometry and expressed in µC.

#### Availability of data and materials

The dataset supporting the conclusions of this article is available in the GEO repository, [GSE75553 in <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE75553>].



## **Competing interests**

The authors declare that they have no competing interests.

## **Authors' contributions**

FA and HF wrote the paper. SMD and AT participated in its compilation. GH, EB, and MBM quantified the amino acids, sugars and organic acids. HP compiled the library and performed the sequencing. MR undertook the demultiplexing and quality control of the reads obtained. MNM cared for the strain. SP performed quantitative PCR. AN performed the statistical analyses. HF mapped the RNA-Seq data and reconstructed the pathways. AM, VA and YLL supervised the work.

## **Acknowledgements**

We are very grateful to the ABIMS platform (CNRS Roscoff, France) for providing scientific computing environment using Galaxy, and assistance with the RNA-seq analysis. This work was *supported by the CAPES-COFECUB French-Brazilian Cooperation Programme*. We also wish to thank the Institut National de la *Recherche Agronomique (INRA, UMR 1253, STLO, Rennes, France)* for its financial and practical support. Our thanks also go to Vicky Hawken for her English revision of the paper.

## Bibliography

1. Isawa K, Hojo K, Yoda N, Kamiyama T, Makino S, Saito M, Sugano H, Mizoguchi C, Kurama S, Shibasaki M, Endo N, Sato Y: **Isolation and identification of a new bifidogenic growth stimulator produced by *Propionibacterium freudenreichii* ET-3.** *Biosci Biotechnol Biochem* 2002, **66**:679–681.
2. Jan G, Rouault A, Maubois JL: **Acid stress susceptibility and acid adaptation of *Propionibacterium freudenreichii* subsp *shermanii*.** *Lait* 2000, **80**:325–336.
3. Okada Y, Tsuzuki Y, Narimatsu K, Sato H, Ueda T, Hozumi H, Sato S, Hokari R, Kurihara C, Komoto S, Watanabe C, Tomita K, Kawaguchi A, Nagao S, Miura S: **1,4-Dihydroxy-2-naphthoic acid from *Propionibacterium freudenreichii* reduces inflammation in interleukin-10-deficient mice with colitis by suppressing macrophage-derived proinflammatory cytokines.** *J Leukoc Biol* 2013, **94**:473–480.
4. Le Marechal C, Peton V, Ple C, Vroland C, Jardin J, Briard-Bion V, Durant G, Chuat V, Loux V, Foligne B, Deutsch S-M, Falentin H, Jan G: **Surface proteins of *Propionibacterium freudenreichii* are involved in its anti-inflammatory properties.** *J Proteomics* 2015, **113**:447–461.
5. Loux V, Mariadassou M, Almeida S, Chiapello H, Hammani A, Buratti J, Gendrault A, Barbe V, Aury J-M, Deutsch S-M, Parayre S, Madec M-N, Chuat V, Jan G, Peterlongo P, Azevedo V, Le Loir Y, Falentin H: **Mutations and genomic islands can explain the strain dependency of sugar utilization in 21 strains of *Propionibacterium freudenreichii*.** *BMC Genomics* 2015, **16**:296.
6. Dalmasso M, Aubert J, Briard-Bion V, Chuat V, Deutsch S-M, Even S, Falentin H, Jan G, Jardin J, Maillard M-B, Parayre S, Piot M, Tanskanen J, Thierry A: **A temporal -omic study of *Propionibacterium freudenreichii* CIRM-BIA1T adaptation strategies in conditions mimicking cheese ripening in the cold.** *PLoS ONE* 2012, **7**:e29083.
7. Benz M, Schink B, Brune A: **Humic acid reduction by *Propionibacterium freudenreichii* and other fermenting bacteria.** *Appl Environ Microbiol* 1998, **64**:4507–4512.
8. de Freitas R, Madec M-N, Chuat V, Maillard M-B, Mukdsi MCA, Falentin H, de Carvalho AF, Valence F, Thierry A: **New insights about phenotypic heterogeneity within *Propionibacterium freudenreichii* argue against its division into subspecies.** *Dairy Sci Technol* 2015, **95**:465–477.
9. Falentin H, Deutsch S-M, Jan G, Loux V, Thierry A, Parayre S, Maillard M-B, Dherbécourt J, Cousin FJ, Jardin J, Siguier P, Couloux A, Barbe V, Vacherie B, Wincker P, Gibrat J-F, Gaillardin C, Lortal S: **The complete genome of *Propionibacterium freudenreichii* CIRM-BIA1T, a hardy Actinobacterium with food and probiotic applications.** *PLoS ONE* 2010, **5**:e11748.
10. Saraoui T, Parayre S, Guernec G, Loux V, Montfort J, Le Cam A, Boudry G, Jan G, Falentin H: **A unique in vivo experimental approach reveals metabolic adaptation of the probiotic *Propionibacterium freudenreichii* to the colon environment.** *BMC Genomics* 2013, **14**:911.

11. Dalmasso M, Aubert J, Even S, Falentin H, Maillard M-B, Parayre S, Loux V, Tanskanen J, Thierry A: **Accumulation of intracellular glycogen and trehalose by Propionibacterium freudenreichii under conditions mimicking cheese ripening in the cold.** *Appl Environ Microbiol* 2012, **78**:6357–6364.
12. Leverrier P, Dimova D, Pichereau V, Auffray Y, Boyaval P, Jan G: **Susceptibility and adaptive response to bile salts in Propionibacterium freudenreichii: physiological and proteomic analysis.** *Appl Environ Microbiol* 2003, **69**:3809–3818.
13. Leverrier P, Fremont Y, Rouault A, Boyaval P, Jan G: **In vitro tolerance to digestive stresses of propionibacteria: influence of food matrices.** *Food Microbiol* 2005, **22**:11–18.
14. Leverrier P, Vissers JP, Rouault A, Boyaval P, Jan G: **Mass spectrometry proteomic analysis of stress adaptation reveals both common and distinct response pathways in Propionibacterium freudenreichii.** *ArchMicrobiol* 2004, **181**:215–230.
15. Aburjaile FF, Madec M-N, Parayre S, Miyoshi A, Azevedo VA, Le Loir Y, Falentin H: **The long-term survival of Propionibacterium freudenreichii in a context of nutrient shortage.** *J Appl Microbiol* 2015, **120**:432-440.
16. Piveteau P: **Metabolism of lactate and sugars by dairy propionibacteria: A review.** *Lait* 1999, **79**:23–41.
17. Jacobsen I, Hennig-Pauka I, Baltes N, Trost M, Gerlach G-F: **Enzymes Involved in Anaerobic Respiration Appear To Play a Role in Actinobacillus pleuropneumoniae Virulence.** *Infect Immun* 2004, **73**:226–234.
18. Deutsch SM, Bivic PL, Herve C, Madec MN, LaPointe G, Jan G, Loir YL, Falentin H: **Correlation of the capsular phenotype in Propionibacterium freudenreichii with the level of expression of gtf, a unique polysaccharide synthase-encoding gene.** *Appl Environ Microbiol* 2010, **76**:2740–2746.
19. Dols-Lafargue M, Lee HY, Le Marrec C, Heyraud A, Chambat G, Lonvaud-Funel A: **Characterization of gtf, a glucosyltransferase gene in the genomes of Pediococcus parvulus and Oenococcus oeni, two bacterial species commonly found in wine.** *Appl Environ Microbiol* 2008, **74**:4079–4090.
20. Loux V, Mariadassou M, Almeida S, Chiapello H, Hammani A, Buratti J, Gendraud A, Barbe V, Aury J-M, Deutsch S-M, Parayre S, Madec M-N, Chuat V, Jan G, Peterlongo P, Azevedo V, Le Loir Y, Falentin H: **Mutations and genomic islands can explain the strain dependency of sugar utilization in 21 strains of Propionibacterium freudenreichii.** *BMC Genomics* 2015, **16**.
21. Gagnaire V, Jardin J, Rabah H, Briard-Bion V, Jan G: **Emmental cheese environment enhances Propionibacterium freudenreichii stress tolerance.** *Plos One* 2015, **10**:e0135780.
22. Crow VL: **Metabolism of aspartate by Propionibacterium-freudenreichii subspecies shermanii - effect on lactate fermentation.** *Appl Environ Microbiol* 1986, **52**:359–365.

23. Blasco L, Kahala M, Tupasela T, Joutsjoki V: **Determination of aspartase activity in dairy Propionibacterium strains.** *Fems Microbiol Lett* 2011, **321**:10–13.
24. EFSA Panel on Biological Hazards (BIOHAZ): **Scientific opinion on scientific opinion on risk based control of biogenic amine formation in fermented foods.** *EFSA J* 2011, **9**:2393.
25. Cotter PD, Gahan CGM, Hill C: **A glutamate decarboxylase system protects Listeria monocytogenes in gastric fluid.** *Mol Microbiol* 2001, **40**:465–475.
26. Kylma AK, Jokela J, Leisola M: **Effect of glycine betaine on osmoadaptation of Propionibacterium acidipropionici cultivated in elevated osmolarities.** *Appl Microbiol Biotechnol* 2000, **54**:705–710.
27. Fox PF, O'Connor TP, McSweeney PLH: **Cheese: physical, biochemical, and nutritional aspects.** *Adv Food Nutr Res* 1996, **39**:163–328.
28. Ishige T, Krause M, Bott M, Wendisch VF, Sahm H: **The phosphate starvation stimulon of Corynebacterium glutamicum determined by DNA microarray analyses.** *J Bacteriol* 2003, **185**:4519–4529.
29. Grillo-Puertas M, Villegas JM, Rintoul MR, Rapisarda VA: **Polyphosphate degradation in stationary phase triggers biofilm formation via luxS quorum sensing system in Escherichia coli.** *PLoS ONE* 2012, **7**:e50368.
30. Rifat D, Bishai WR, Karakousis PC: **Phosphate depletion: a novel trigger for Mycobacterium tuberculosis persistence.** *Infect Dis* 2009, **200**:1126–1135.
31. Jan G, Leverrier P, Pichereau V, Boyaval P: **Changes in protein synthesis and morphology during acid adaptation of Propionibacterium freudenreichii.** *Appl Environ Microbiol* 2001, **67**:2029–2036.
32. Claverys J-P, Prudhomme M, Martin B: **Induction of competence regulons as a general response to stress in Gram-Positive bacteria.** *Annu Rev Microbiol* 2006, **60**:451–475.
33. Vos M: **Why do bacteria engage in homologous recombination?** *Trends Microbiol* 2009, **17**:226–232.
34. Dalmasso M, Nicolas P, Falentin H, Valence F, Tanskanen J, Jatila H, Salusjärvi T, Thierry A: **Multilocus sequence typing of Propionibacterium freudenreichii.** *Int J Food Microbiol* 2011, **145**:113–120.
35. Zhang Y, Liu J-Z, Huang J-S, Mao Z-W: **Genome shuffling of Propionibacterium shermanii for improving vitamin B12 production and comparative proteome analysis.** *J Biotechnol* 2010, **148**:139–143.
36. Malik AC, Reinbold GW, Vedamuthu ER: **Evaluation of the taxonomy of the Propionibacterium.** *Canadian Journal of Microbiology* 1968:1185–1191.
37. Baron F, Cochet MF, Ablain W, Grosset N, Madec MN, Gonnet F, Jan S, Gautier M: **Rapid and cost-effective method for microorganism enumeration based on**

**miniaturization of the conventional plate-counting technique.** *Lait* 2006, **86**:251–257.

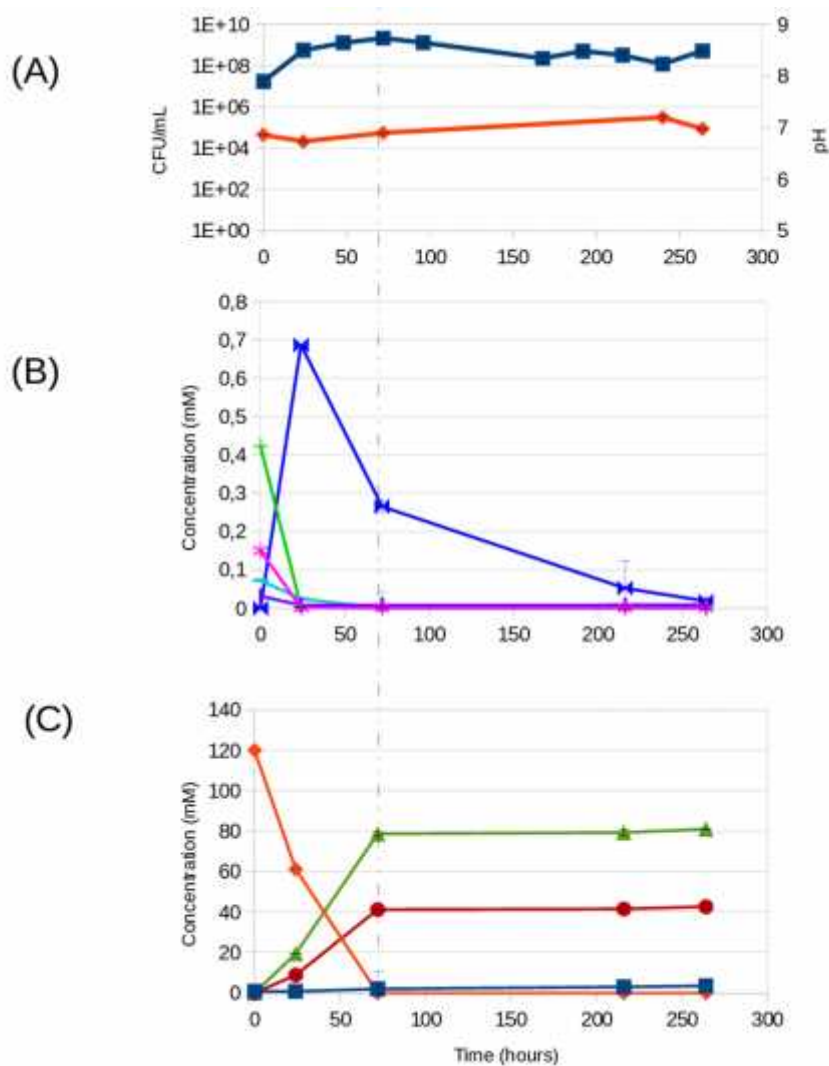
38. Langmead B, Trapnell C, Pop M, Salzberg SL: **Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.** *Genome Biol* 2009, **10**:R25.

39. Anders S, Pyl PT, Huber W: **HTSeq--a Python framework to work with high-throughput sequencing data.** *Bioinformatics* 2015, **31**:166–169.

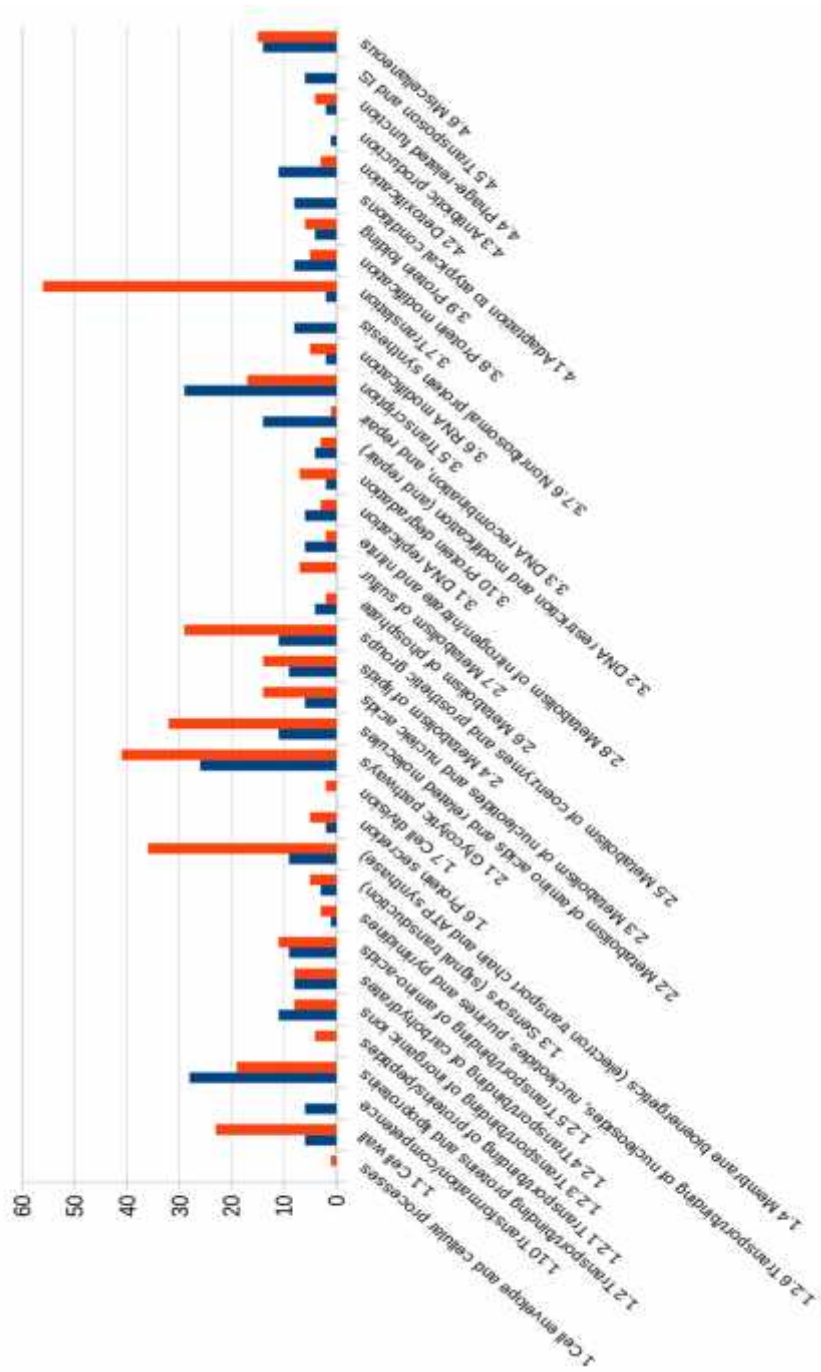
40. Love MI, Huber W, Anders S: **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.** *Genome Biol* 2014,**15**:550.

41. Robinson MD, McCarthy DJ, Smyth GK: **edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.** *Bioinformatics* 2010, **26**:139–140.

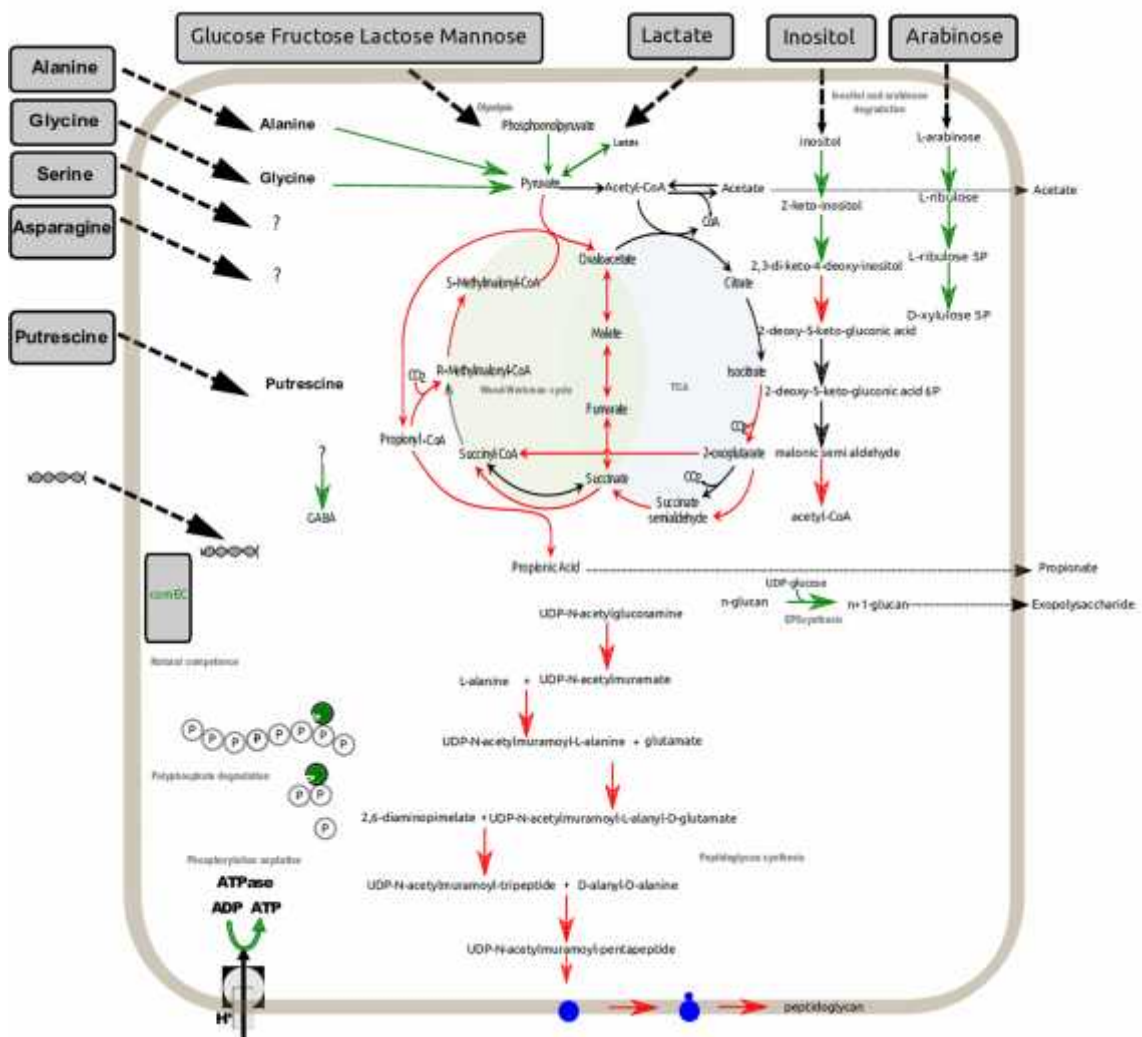
42. Luo W, Friedman MS, Shedden K, Hankenson KD, Woolf PJ: **GAGE: generally applicable gene set enrichment for pathway analysis.** *BMC Bioinformatics* 2009, **10**:161.



**Figure 1:** DO, CFU, pH, sugar utilization and organic acid consumption and production during the growth and long term survival over 11 days of CIRM-BIA138 in YEL medium. (A) Bacterial enumeration (CFU/mL), and pH (B) Quantification of sugars (glucose: pink stars, mannose: purple bold pipes, fructose: light blue dashes, lactose: green pipes) for conversion to pyruvate (blue bow tie). (C) Quantification of acids (succinate: dark blue square, lactate: orange diamond, acetate: red circle, propionate: green triangle). The vertical dotted line indicates entry into stationary phase (3d post-inoculation).

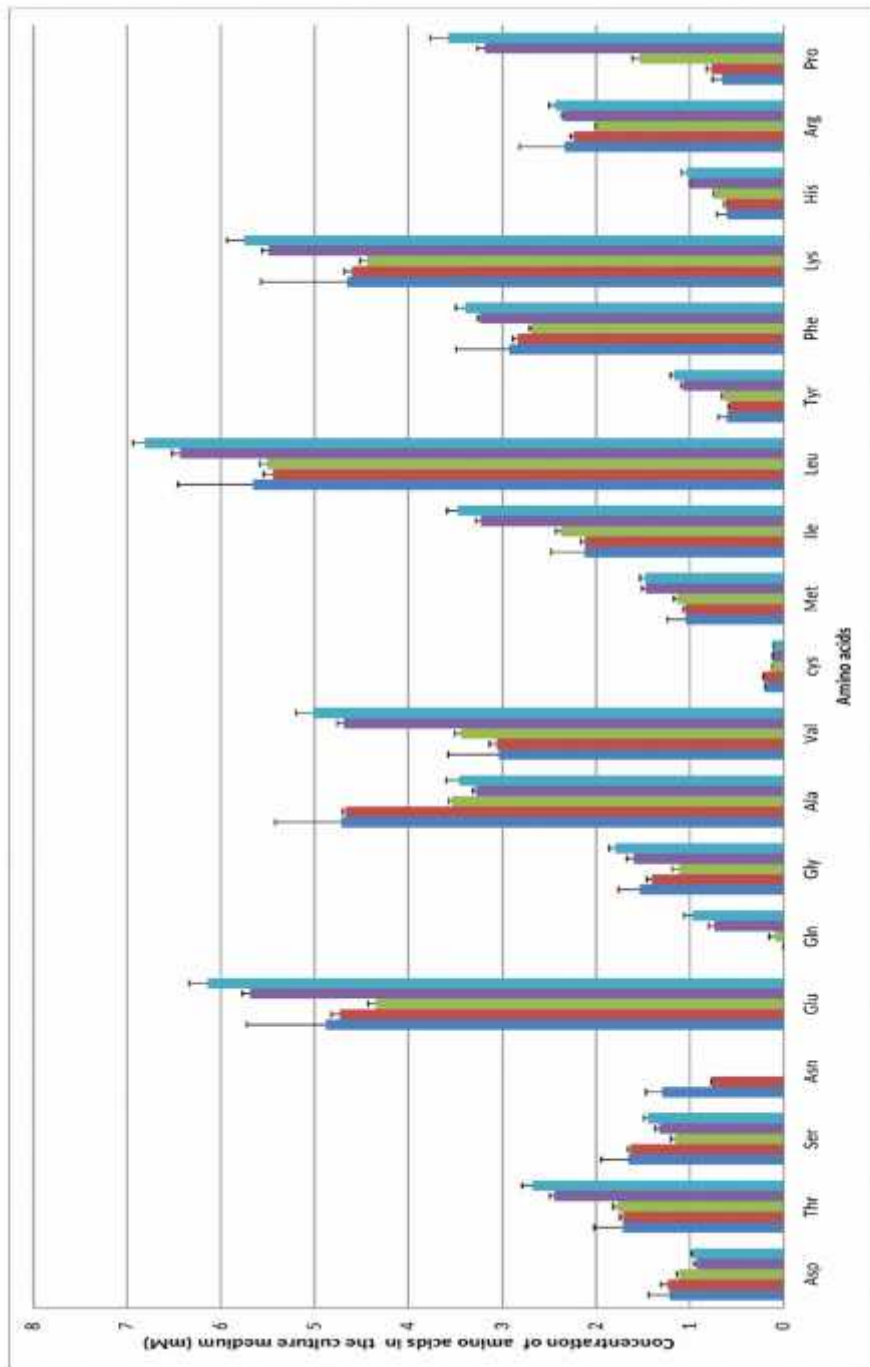


**Figure 2:** Number of induced and repressed genes in the stationary phase (3d post-inoculation) compared to the exponential phase (1d post-inoculation), classified by metabolic function. In blue and orange: induced and repressed genes, respectively. Differential comparison of groups (each containing three repetitions) were performed gene by gene using a modified t-test. Genes were declared as being differentially expressed (DE) with a P value 0.005.



**Figure 3:** Schematic view of the metabolism at the entry into stationary phase compared to exponential phase. Induced / repressed / 'non differentially expressed' genes are symbolized with green / red / black arrows respectively. (P value = 0.005).





**Figure 4:** Concentrations of free amino acids in the medium (mM) during growth and stationary phases. In dark blue at the time of inoculation, in red at 1-day post inoculation, in green at 3-day post inoculation, in purple at 9-day post-inoculation, in light blue at 11-day post-inoculation. Significant consumptions discussed in the text (serine, asparagine, glycine, alanine) correspond to the difference in concentration at 1d and 3d post inoculation (t-test, Pvalue <0.05).

**Table S1** : Induced genes in stationary phase compared to exponential phase.

Name	Descript	Function	FC	log2Fold Change	pvalue	padj
mraZ	Protein mraZ	1.1 Cell wall	1,839	0,879	0,01458547	0,03763326
gtfF	Polysaccharide synthase	1.1 Cell wall	6,169	2,625	0,0011783	0,00431847
slpE	Surface protein with SLH domain	1.1 Cell wall	12,903	3,69	2,7762E-07	2,965E-06
PFCIRM13 8_06855	Hypothetical outer membrane protein	1.1 Cell wall	4,992	2,32	0,00269909	0,00896146
gtfD	Glycosyl transferase, family 2	1.1 Cell wall	2,274	1,185	1,7102E-05	0,00011405
PFCIRM13 8_07670	nucleoside-diphosphate-sugar epimerases	1.1 Cell wall	2,194	1,134	2,3011E-05	0,00014509
PFCIRM13 8_00630	DNA transfert protein	1.10 Transformation/competence	14,752	3,883	2,1136E-05	0,00013617
dprA	DNA processing / uptake protein	1.10 Transformation/competence	27,362	4,774	1,0898E-07	1,2959E-06
PFCIRM13 8_07015	Mobilisation protein	1.10 Transformation/competence	49,447	5,628	4,7645E-40	5,2124E-37
PFCIRM13 8_07020	Mobilisation protein	1.10 Transformation/competence	10,855	3,44	6,5035E-22	1,423E-19
PFCIRM13 8_07570	Conjugative relaxase domain protein, TrwC/Tral family	1.10 Transformation/competence	24,025	4,586	2,7937E-06	2,4065E-05
comEC	ComE operon protein 3	1.10 Transformation/competence	17,39	4,12	1,2834E-06	1,2E-05
PFCIRM13 8_00130	Permease	1.2 Transport/binding proteins and	20,821	4,38	7,3817E-07	7,1465E-06

		lipoproteins				
PFCIRM13 8_00900	ABC transporter (Permease)	1.2 Transport/binding proteins and lipoproteins	12,247	3,614	7,9844E-05	0,0004261
sufS	Cysteine desulphurases, SufS	1.2 Transport/binding proteins and lipoproteins	1,956	0,968	0,00877986	0,02440956
sufC	ABC-type transport system involved in Fe-S cluster assembly, ATPase component, SufC	1.2 Transport/binding proteins and lipoproteins	1,926	0,945	0,01558108	0,039707
sufD	FeS assembly protein SufD	1.2 Transport/binding proteins and lipoproteins	2,114	1,08	0,00534908	0,01603257
phnL	ATP binding protein of ABC transporter	1.2 Transport/binding proteins and lipoproteins	15,223	3,928	1,0113E-06	9,5378E-06
PFCIRM13 8_02385	permease protein	1.2 Transport/binding proteins and lipoproteins	50,478	5,658	1,6649E-06	1,5115E-05
PFCIRM13 8_02395	ABC transporter related precursor	1.2 Transport/binding proteins and lipoproteins	12,372	3,629	0,00015522	0,00075138
PFCIRM13 8_03530	Binding-protein-dependent transport systems inner membrane component	1.2 Transport/binding proteins and lipoproteins	2,246	1,168	0,00035727	0,00156031
yceJ	Transporter protein	1.2 Transport/binding proteins and lipoproteins	16,874	4,077	2,9259E-05	0,00017882
PFCIRM13 8_05125	Binding-protein-dependent transport systems inner membrane component	1.2 Transport/binding proteins and lipoproteins	10,809	3,434	0,00017228	0,00081767
PFCIRM13 8_05130	Binding-protein-dependent transport systems inner membrane component	1.2 Transport/binding proteins and lipoproteins	13,72	3,778	3,2365E-05	0,00019401
PFCIRM13 8_05560	Transporter protein	1.2 Transport/binding proteins and lipoproteins	6,834	2,773	0,00103039	0,00382765

		lipoproteins				
PFCIRM13 8_06235	Phosphocarrier protein, HPr family 8.A.8.1.1	1.2 Transport/binding proteins and lipoproteins	1,736	0,796	0,02059267	0,04945857
PFCIRM13 8_06470	ABC-type transporter, permease component	1.2 Transport/binding proteins and lipoproteins	3,067	1,617	0,00136804	0,00490701
PFCIRM13 8_06475	BP protein of siderophore ABC transporter (ISVH:undef)	1.2 Transport/binding proteins and lipoproteins	3,578	1,839	0,00071747	0,00280327
PFCIRM13 8_07645	transmembrane efflux protein	1.2 Transport/binding proteins and lipoproteins	19,201	4,263	1,4006E-05	9,6067E-05
PFCIRM13 8_08270	ABC transporter ATP-binding protein	1.2 Transport/binding proteins and lipoproteins	8,213	3,038	0,00100895	0,00375439
ImrB	Permease (Lincomycin resistance protein )	1.2 Transport/binding proteins and lipoproteins	13,929	3,8	0,00012318	0,00061956
PFCIRM13 8_08660	secretory protein	1.2 Transport/binding proteins and lipoproteins	10,949	3,453	0,00020008	0,00093742
PFCIRM13 8_09385	Arsenite efflux pump ACR3 and related permeases (Arsenite efflux pump ACR3 or related permease)	1.2 Transport/binding proteins and lipoproteins	3,001	1,586	0,0013183	0,00477556
PFCIRM13 8_09450	major facilitator superfamily drug:H <sup>+</sup> antiporter 2.A.1.3.x	1.2 Transport/binding proteins and lipoproteins	6,824	2,771	2,1795E-19	2,3844E-17
cynX	Cyanate MSF transporter	1.2 Transport/binding proteins and lipoproteins	3,305	1,725	0,00830715	0,02339258
pimH	Integral membrane efflux protein	1.2 Transport/binding proteins and lipoproteins	4,234	2,082	0,00105323	0,00390588
hflC2	Stomatin/prohibitin	1.2 Transport/binding proteins and	3,522	1,816	3,0476E-09	4,9394E-08

		lipoproteins				
PFCIRM13 8_10665	ABC transporter ATP binding	1.2 Transport/binding proteins and lipoproteins	2,084	1,06	0,00756706	0,02162374
PFCIRM13 8_10695	IM-ABC protein of drug, protein or lipid ABC transporter (DPL:LLP)	1.2 Transport/binding proteins and lipoproteins	3,161	1,66	0,00117113	0,00430662
PFCIRM13 8_12260	transporter, MFS superfamily protein	1.2 Transport/binding proteins and lipoproteins	11,237	3,49	7,5623E-05	0,00040654
PFCIRM13 8_05030	Cobalt permease	1.2.3 Transport/binding of inorganic ions	2,753	1,461	0,01803761	0,04454435
PFCIRM13 8_05035	ABC-type transporter, ATPase component	1.2.3 Transport/binding of inorganic ions	2,48	1,31	0,00253841	0,00849243
mntH	Manganese transport protein	1.2.3 Transport/binding of inorganic ions	7,261	2,86	8,9333E-05	0,00047203
PFCIRM13 8_08475	nitrate transporter	1.2.3 Transport/binding of inorganic ions	1,88	0,911	0,00799075	0,02261805
ssuC	permease protein of nitrate ABC transporter	1.2.3 Transport/binding of inorganic ions	4,413	2,142	0,00047049	0,00196081
PFCIRM13 8_08620	Heavy metal transport/detoxification protein	1.2.3 Transport/binding of inorganic ions	3,759	1,91	6,4909E-07	6,4263E-06
PFCIRM13 8_08625	Cation-transporting ATPase	1.2.3 Transport/binding of inorganic ions	4,248	2,087	5,1528E-08	6,5548E-07
PFCIRM13 8_09635	divalent anion:Na <sup>+</sup> symporter 2.A.47.3.x	1.2.3 Transport/binding of inorganic ions	2,399	1,262	5,8332E-05	0,00032393
feuS	binding protein of iron ABC transporter	1.2.3 Transport/binding of inorganic ions	2,805	1,488	0,00559075	0,01668836
feuC	inner membrane protein of iron ABC transporter	1.2.3 Transport/binding of inorganic ions	2,468	1,303	0,01488094	0,03826028
PFCIRM13 8_11450	Sodium:sulfate symporter transmembrane region	1.2.3 Transport/binding of inorganic ions	4,306	2,106	0,0059313	0,01753638
PFCIRM13 8_04135	permease of the major facilitator superfamily	1.2.4 Transport/binding of carbohydrates	3,582	1,841	2,0847E-06	1,8467E-05
PFCIRM13	Membrane permease (sugar transporter)	1.2.4 Transport/binding	9,736	3,283	6,3679E-21	8,7081E-19

8_04640		of carbohydrates				
ptsG/ptsM	PTS system mannose-specific EIIBCA component (EIIBCA-Man) (EII-Man/EIII-Man)	1.2.4 Transport/binding of carbohydrates	2,025	1,018	0,00072156	0,00281422
ptsI	PTS enzyme I	1.2.4 Transport/binding of carbohydrates	2,309	1,207	0,0001149	0,00058876
yaaU	metabolite transport protein YaaU	1.2.4 Transport/binding of carbohydrates	4,656	2,219	0,00050151	0,00205875
gudP	D-galactonate transporter	1.2.4 Transport/binding of carbohydrates	10,093	3,335	5,6868E-05	0,00031741
ioIT1	ioIT1 (Major myo-inositol transporter ioIT1)	1.2.4 Transport/binding of carbohydrates	11,044	3,465	1,9084E-05	0,00012577
araH2	inner membrane protein of ribose/xylose/arabinose/galactoside ABC transporter	1.2.4 Transport/binding of carbohydrates	3,838	1,94	0,0083652	0,02352578
PFCIRM13 8_00815	Glutamate/gamma-aminobutyrate anti-porter (membrane protein)	1.2.5 Transport/binding of amino-acids	8,595	3,104	3,0355E-05	0,0001833
PFCIRM13 8_04335	polar amino acid ABC transporter, ATP-binding protein	1.2.5 Transport/binding of amino-acids	33,689	5,074	8,5244E-07	8,2165E-06
PFCIRM13 8_04340	polar amino acid ABC transporter, binding protein component	1.2.5 Transport/binding of amino-acids	10,43	3,383	0,00016304	0,00078057
PFCIRM13 8_04345	Polar amino acid ABC transporter, inner membrane subunit	1.2.5 Transport/binding of amino-acids	14,928	3,9	3,0824E-05	0,00018528
PFCIRM13 8_04350	polar amino acid ABC transporter, inner membrane protein	1.2.5 Transport/binding of amino-acids	16,067	4,006	3,041E-05	0,0001833
lysE	Lysine exporter protein, membran protein	1.2.5 Transport/binding of amino-acids	4,953	2,308	0,01065826	0,02875496
PFCIRM13 8_07335	polar amino acid ABC transporter, binding protein	1.2.5 Transport/binding of amino-acids	3,294	1,72	0,00149863	0,00531441
PFCIRM13 8_07665	Amino acid permease-associated region	1.2.5 Transport/binding of amino-acids	2,317	1,212	0,00040015	0,00171335
PFCIRM13 8_10015	Putrescine importer	1.2.5 Transport/binding of amino-acids	3,541	1,824	1,1329E-09	2,0486E-08
cytX	Hydroxymethylpyrimidine transporter CytX	1.2.6 Transport/binding of nucleosides, nucleotides, purines and pyrimidines	3,892	1,961	0,01324767	0,03458939

PFCIRM13 8_08430	sensor kinase - Membrane protein	1.3 Sensors (signal transduction)	5,331	2,414	2,2078E-05	0,00014042
PFCIRM13 8_10375	Sensory transduction histidine kinase	1.3 Sensors (signal transduction)	1,555	0,637	0,00918295	0,02536905
PFCIRM13 8_11140	Two-component system sensor kinase	1.3 Sensors (signal transduction)	1,953	0,966	0,00048526	0,00201087
PFCIRM13 8_01505	Dioxygenase	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	1,987	0,99	0,00871079	0,02434126
PFCIRM13 8_07220	Cytochrome c biogenesis protein transmembrane region	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	13,347	3,738	9,8363E-05	0,00051243
dmsC	Anaerobic dimethyl sulfoxide reductase chain C	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	13,75	3,781	0,000144	0,00070486
dmsB	Anaerobic dimethyl sulfoxide reductase, chain B	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	4,764	2,252	0,00987035	0,02696171
PFCIRM13 8_08930	NAD(P)H-dependent FMN reductase	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	2,962	1,567	3,4375E-06	2,8707E-05
mer	Coenzyme F420-dependent N5,N10-methylene tetrahydromethanopterin reductase	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	3,124	1,643	2,7322E-07	2,9448E-06
PFCIRM13 8_09375	alkylglycerone-phosphate synthase (FAD linked oxidase-like protein)	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	10,084	3,334	2,5391E-14	9,9207E-13
PFCIRM13 8_10905	NADPH:quinone reductase and related Zn-dependent oxidoreductases	1.4 Membrane bioenergetics (electron	1,718	0,781	0,01159156	0,030892

		transport chain and ATP synthase)				
PFCIRM13 8_11130	NADPH:quinone reductase related Zn-dependent oxidoreductase	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	2,623	1,391	2,2365E-05	0,00014184
PFCIRM13 8_06860	Peptidase C60, sortase A and B precursor	1.6 Protein secretion	22,196	4,472	1,2374E-05	8,7054E-05
PFCIRM13 8_07165	Tat pathway signal sequence domain protein	1.6 Protein secretion	15,425	3,947	1,3403E-05	9,2323E-05
PFCIRM13 8_00105	protein of unknown function	10.2 Protein of unknown function similar to proteins from other organisms	2,054	1,039	0,01220287	0,03224623
PFCIRM13 8_05230	Zinc-containing alcohol dehydrogenase	2 Intermediary metabolism	15,475	3,952	3,329E-23	9,1049E-21
gntK	gluconokinase	2.1.1 Specific carbohydrate metabolic pathway	2,572	1,363	1,1071E-05	7,9608E-05
galK1	Galactokinase	2.1.1 Specific carbohydrate metabolic pathway	1,999	1	0,01117885	0,03001144
PFCIRM13 8_00835	Inositol 2-dehydrogenase	2.1.1 Specific carbohydrate metabolic pathway	7,458	2,899	0,00060532	0,00242127
PFCIRM13 8_00840	sugar (xylose) phosphate isomerase/epimerase	2.1.1 Specific carbohydrate metabolic pathway	14,326	3,841	0,00016245	0,00077946
iolH	iolH (Myo-inositol catabolism IolH protein) (isomerase:epimerase)	2.1.1 Specific carbohydrate metabolic pathway	23,005	4,524	3,6816E-05	0,00021713
glgA	glycosyltransferase (glycogen synthase)	2.1.1 Specific carbohydrate metabolic pathway	2,264	1,179	4,0563E-05	0,00023604
iolT2	iolT2 (myo-inositol transporter iolT2)	2.1.1 Specific carbohydrate metabolic	3,219	1,686	0,00447323	0,01370789



		pathway				
iol	myo-inositol 2-dehydrogenase	2.1.1 Specific carbohydrate metabolic pathway	5,559	2,475	0,00085922	0,00326954
PFCIRM13 8_03930	Glycoside hydrolase	2.1.1 Specific carbohydrate metabolic pathway	2,345	1,229	0,0184331	0,04526557
adh	alcohol dehydrogenase	2.1.1 Specific carbohydrate metabolic pathway	10,194	3,35	0,00036667	0,001595
PFCIRM13 8_04360	hydroxyacid dehydrogenase	2.1.1 Specific carbohydrate metabolic pathway	14,233	3,831	7,0972E-05	0,00038627
aldB	Lactaldehyde dehydrogenase	2.1.1 Specific carbohydrate metabolic pathway	2,665	1,414	1,13E-05	8,0798E-05
PFCIRM13 8_06500	Beta-glucosidase	2.1.1 Specific carbohydrate metabolic pathway	1,621	0,696	0,01711601	0,04270219
PFCIRM13 8_07275	hexulose-6-phosphate isomerase	2.1.1 Specific carbohydrate metabolic pathway	3,863	1,95	0,00708094	0,02033216
malQ2	4-alpha-glucanotransferase (Amylomaltase) (Disproportionating enzyme) (D-enzyme)	2.1.1 Specific carbohydrate metabolic pathway	1,529	0,613	0,01739521	0,04319038
galE3	UDP-glucose 4-epimerase	2.1.1 Specific carbohydrate metabolic pathway	11,493	3,523	0,00010621	0,0005494
iolG2	iolG2 (Myo-inositol catabolism IolG2 protein) (Inositol 2-dehydrogenase)	2.1.1 Specific carbohydrate metabolic pathway	2,114	1,08	0,0116371	0,03097563
iolE3	Myo-inositol catabolism protein	2.1.1 Specific carbohydrate metabolic pathway	2,26	1,176	0,00100895	0,00375439
araD1	L-ribulose-5-phosphate 4-epimerase	2.1.1 Specific carbohydrate metabolic	2,869	1,52	0,007871	0,02233689

		pathway				
araA	L-arabinose isomerase	2.1.1 Specific carbohydrate metabolic pathway	2,187	1,129	0,01676094	0,04215279
fba2	Fructose-bisphosphate aldolase class I	2.1.2 Main glycolytic pathways	3,651	1,868	5,6726E-09	8,6795E-08
rpiB3	Ribose-5-phosphate isomerase 3	2.1.2 Main glycolytic pathways	2,879	1,526	3,1493E-06	2,6503E-05
acyP	Putative acylphosphatase	2.1.2 Main glycolytic pathways	5,76	2,526	0,0016336	0,00574651
araB	L-ribulokinase	2.1.2 Main glycolytic pathways	2,158	1,11	1,2915E-05	8,999E-05
eno2	Enolase 2	2.1.2 Main glycolytic pathways	3,442	1,783	1,2064E-08	1,7598E-07
PFCIRM13 8_07350	hydroxypyruvate isomerase	2.1.4 Substrate-specific entries to carbohydrate metabolic pathway	15,014	3,908	1,0201E-05	7,4646E-05
leuC	3-isopropylmalate dehydratase large subunit (Isopropylmalate isomerase) (Alpha-IPM isomerase) (IPMI)	2.2 Metabolism of amino acids and related molecules	2,413	1,271	0,00049172	0,00202233
gadB	Glutamate decarboxylase	2.2 Metabolism of amino acids and related molecules	4,997	2,321	2,3166E-06	2,0194E-05
aroF	Chorismate synthase (5-enolpyruvylshikimate-3-phosphate phospholyase)	2.2 Metabolism of amino acids and related molecules	1,569	0,65	0,01874808	0,0457311
metK	S-adenosylmethionine synthetase (Methionine adenosyltransferase) (AdoMet synthetase) (MAT)	2.2 Metabolism of amino acids and related molecules	1,608	0,685	0,02029457	0,04884985
ald	Alanine dehydrogenase	2.2 Metabolism of amino acids and related molecules	2,668	1,416	6,0251E-05	0,0003329
dadA2	DadA, Glycine/D-amino acid oxidases	2.2 Metabolism of amino acids and related molecules	12,994	3,7	4,9033E-06	3,9442E-05

PFCIRM13 8_04365	aminotransferase	2.2 Metabolism of amino acids and related molecules	10,633	3,41	0,00021657	0,00100392
PFCIRM13 8_05235	Carbamoylphosphate synthase large subunit CarB	2.2 Metabolism of amino acids and related molecules	4,228	2,08	7,2777E-06	5,5677E-05
aroA	3-phosphoshikimate 1-carboxyvinyltransferase (phosphoenolpyruvate:3-phosphoshikimate5-O-(1-carboxyvinyl)-transferase) (EPSP synthase)	2.2 Metabolism of amino acids and related molecules	1,672	0,742	0,00276818	0,00912165
dapA1	Dihydrodipicolinate synthetase	2.2 Metabolism of amino acids and related molecules	38,828	5,279	2,6405E-07	2,8602E-06
leuA1	2-isopropylmalate synthase	2.2 Metabolism of amino acids and related molecules	1,676	0,745	0,00347887	0,01090968
ppx	Ppx/GppA phosphatase family	2.3 Metabolism of nucleotides and nucleic acids	1,62	0,696	0,02002399	0,04827567
nrdR	Transcriptional repressor nrdR	2.3 Metabolism of nucleotides and nucleic acids	1,728	0,789	0,01925179	0,04669946
thiD1	Phosphomethylpyrimidine kinase/ hydroxymethylpyrimidine phosphokinase	2.3 Metabolism of nucleotides and nucleic acids	4,07	2,025	0,0111065	0,02989055
dgt	Deoxyguanosinetriphosphate triphosphohydrolase (dGTPase)	2.3 Metabolism of nucleotides and nucleic acids	1,822	0,866	0,00312999	0,00996859
prt	phosphoribosyltransferase	2.3 Metabolism of nucleotides and nucleic acids	6,778	2,761	0,00026501	0,00119308
pcaB2	Adenylosuccinate lyase	2.3 Metabolism of nucleotides and nucleic acids	11,911	3,574	0,00021726	0,00100499
pf1655	Carboxylic ester hydrolase	2.4 Metabolism of lipids	1,511	0,595	0,01741038	0,04319038
pf169	Carboxylic ester hydrolase	2.4 Metabolism of lipids	2,308	1,207	0,00380039	0,01182825

PFCIRM13 8_03645	Esterase family protein	2.4 Metabolism of lipids	65,916	6,043	4,4128E-11	1,0495E-09
caiA	Crotonobetainyl-CoA dehydrogenase (Crotonobetainyl-CoA reductase)	2.4 Metabolism of lipids	15,443	3,949	1,006E-05	7,386E-05
caiB	Crotonobetainyl-CoA:carnitine CoA-transferase	2.4 Metabolism of lipids	3,281	1,714	0,00683535	0,0197566
caiC	Crotonobetaine/carnitine-CoA ligase	2.4 Metabolism of lipids	12,384	3,63	4,7644E-06	3,8467E-05
uppS	Undecaprenyl pyrophosphate synthase (di-trans,poly-cis-decaprenylcistransferase)	2.4 Metabolism of lipids	1,789	0,839	0,00412185	0,01270227
glxK/garK	Glycerate kinase GlxK/GarK	2.4 Metabolism of lipids	2,794	1,482	2,8322E-08	3,8731E-07
plsC2	PlsC, Phosphate acyltransferases	2.4 Metabolism of lipids	2,069	1,049	0,00527008	0,01581748
chlI	Magnesium chelatase, subunit ChII	2.5 Metabolism of coenzymes and prosthetic groups	8,331	3,058	5,9295E-05	0,00032845
PFCIRM13 8_04810	Isochorismatase hydrolase	2.5 Metabolism of coenzymes and prosthetic groups	1,848	0,886	0,01286699	0,03379709
hemX	Membrane-bound transport protein (HemX protein)	2.5 Metabolism of coenzymes and prosthetic groups	16,735	4,065	2,6276E-05	0,00016287
pdxA2	4-hydroxythreonine-4-phosphate dehydrogenase	2.5 Metabolism of coenzymes and prosthetic groups	2,549	1,35	0,00083498	0,00319352
pncA	pyrazinamidase/nicotinamidase	2.5 Metabolism of coenzymes and prosthetic groups	2,501	1,322	3,4305E-06	2,8707E-05
cbiO3	ABC transporter, ATP-binding protein	2.5 Metabolism of coenzymes and prosthetic groups	5,571	2,478	0,00229833	0,00776041
cbiQ1	ABC-type transporter, permease components (ABC-type cobalt transport system, permease component CbiQ)	2.5 Metabolism of coenzymes and prosthetic groups	9,615	3,265	0,00026681	0,00119871
mog	Molybdopterin biosynthesis Mog protein	2.5 Metabolism of coenzymes and prosthetic groups	2,792	1,481	0,00011522	0,00058903
thiC	thiamine-phosphate diphosphorylase	2.5 Metabolism of	2,361	1,239	0,0156067	0,039707

		coenzymes and prosthetic groups				
PFCIRM13_8_09275	NADH-flavin reductase	2.5 Metabolism of coenzymes and prosthetic groups	1,801	0,849	0,0066313	0,01926864
bioB	Biotin synthase	2.5 Metabolism of coenzymes and prosthetic groups	7,499	2,907	4,262E-27	1,5542E-24
ppx5	Exopolyphosphatase	2.6 Metabolism of phosphate	19,842	4,31	9,256E-06	6,8885E-05
ppa	Inorganic pyrophosphatase	2.6 Metabolism of phosphate	1,579	0,659	0,01239683	0,03267982
PFCIRM13_8_07685	NUDIX hydrolase	2.6 Metabolism of phosphate	3,022	1,595	0,00035576	0,00155682
phoH	Phosphate starvation-inducible protein PhoH-like protein	2.6 Metabolism of phosphate	2,521	1,334	4,7091E-05	0,00027115
ntcA2	Global nitrogen regulator NtcA	2.8 Metabolism of nitrogen/nitrate and nitrite	1,732	0,792	0,00180183	0,00626775
PFCIRM13_8_04265	Nitrate reductase, alpha subunit	2.8 Metabolism of nitrogen/nitrate and nitrite	5,948	2,572	0,00096256	0,00361871
narH	Respiratory nitrate reductase, beta subunit	2.8 Metabolism of nitrogen/nitrate and nitrite	22,415	4,486	3,9715E-07	4,1183E-06
narJ	Nitrate reductase delta subunit	2.8 Metabolism of nitrogen/nitrate and nitrite	22,951	4,52	3,0582E-06	2,5835E-05
narI	Respiratory nitrate reductase, gamma subunit	2.8 Metabolism of nitrogen/nitrate and nitrite	16,346	4,031	7,471E-06	5,6957E-05
glnB	Nitrogen regulatory protein P-II	2.8 Metabolism of nitrogen/nitrate and nitrite	4,518	2,176	0,00150854	0,00534091
dnaE2	DNA polymerase III alpha subunit	3.1 DNA replication	2,304	1,204	0,00804479	0,0227416

priA	Primosomal protein N	3.1 DNA replication	2,979	1,575	5,5968E-09	8,6311E-08
gyrB2	DNA gyrase subunit B	3.1 DNA replication	2,435	1,284	1,9573E-05	0,00012784
hels	Helicase	3.1 DNA replication	2,469	1,304	5,1448E-07	5,2115E-06
PFCIRM13 8_08320	helicase	3.1 DNA replication	2,774	1,472	6,7193E-08	8,2594E-07
dnaB	Replicative DNA helicase	3.1 DNA replication	1,552	0,635	0,01620308	0,04103201
PFCIRM13 8_09485	Membrane protease subunits, stomatin/prohibitin homologs (Membrane protease subunit, stomatin/prohibitin homolog)	3.10 Protein degradation	2,516	1,331	1,4334E-05	9,77E-05
pepC	Aminopeptidase C, Bleomycin hydrolase,	3.10 Protein degradation	1,601	0,679	0,0092136	0,02542164
tagA	3-Methyladenine DNA glycosylase I	3.2 DNA restriction and modification (and repair)	3,915	1,969	0,01295853	0,03397163
sms	DNA repair protein	3.2 DNA restriction and modification (and repair)	1,904	0,929	0,00144176	0,00513773
sbcD	SbcD, DNA repair exonuclease	3.2 DNA restriction and modification (and repair)	2,785	1,478	2,8782E-06	2,4504E-05
PFCIRM13 8_10795	exonuclease	3.2 DNA restriction and modification (and repair)	1,828	0,87	0,01437546	0,03722308
PFCIRM13 8_00325	Recombinase	3.3 DNA recombination, and repair	77,776	6,281	6,6498E-07	6,4955E-06
gin	Resolvase	3.3 DNA recombination, and repair	1,823	0,866	0,01302721	0,03409513
PFCIRM13 8_00650	Recombinase	3.3 DNA recombination, and repair	7,929	2,987	9,3006E-18	7,2678E-16
recG	ATP-dependent DNA helicase RecG	3.3 DNA recombination, and repair	1,74	0,799	0,00307427	0,0098197
dinP	DNA-directed DNA polymerase	3.3 DNA	5,915	2,564	0,0006119	0,00243761

		recombination, and repair				
PFCIRM13_8_01865	Helicase protein	3.3 DNA recombination, and repair	1,806	0,853	0,00091498	0,00345169
recA	RecA (Recombinase A)	3.3 DNA recombination, and repair	2,019	1,014	0,0023503	0,00790909
PFCIRM13_8_03835	Bifunctional DNA primase/polymerase	3.3 DNA recombination, and repair	18,618	4,219	6,2207E-43	1,3611E-39
rmuC	DNA recombination protein	3.3 DNA recombination, and repair	3,851	1,945	2,0143E-10	4,3208E-09
sbcC	ATP-dependent dsDNA exonuclease SbcC	3.3 DNA recombination, and repair	1,712	0,776	0,00302292	0,00971241
alkA	AlkA, 3-methyladenine DNA glycosylase/8-oxoguanine DNA glycosylase	3.3 DNA recombination, and repair	7,574	2,921	0,00061275	0,00243761
PFCIRM13_8_10165	ATP-dependent DNA helicase	3.3 DNA recombination, and repair	2,875	1,524	6,3437E-07	6,3091E-06
urvD	UvrD/REP helicase / ATP-dependent DNA helicase	3.3 DNA recombination, and repair	1,534	0,617	0,0186472	0,04558667
holA	DNA polymerase III, delta subunit	3.3 DNA recombination, and repair	3,066	1,617	1,1811E-07	1,3969E-06
hrdD2	RNA polymerase principal sigma factor HrdD	3.5.1 Transcription initiation	6,278	2,65	3,3471E-10	6,909E-09
hrdD1	RNA polymerase principal sigma factor HrdD	3.5.1 Transcription initiation	8,762	3,131	0,00020197	0,00094223
rpoE	RNA polymerase sigma -70 factor, sigma-E factor	3.5.1 Transcription initiation	1,875	0,907	0,01422102	0,03686681
marR3	Transcriptional regulator, MarR family	3.5.2 Transcription	3,689	1,883	0,01723654	0,04295393

		regulation				
PFCIRM13 8_00405	Transcriptional regulator	3.5.2 Transcription regulation	3,214	1,684	6,4352E-06	4,993E-05
arsR1	Transcriptional regulator, ArsR family protein	3.5.2 Transcription regulation	6,938	2,794	1,0324E-05	7,5048E-05
dtxR	Iron-dependent repressor	3.5.2 Transcription regulation	1,512	0,596	0,01689512	0,0424186
PFCIRM13 8_00890	transcription regulator	3.5.2 Transcription regulation	3,484	1,801	6,3071E-06	4,9285E-05
arsR4	Transcriptional regulator, ArsR family	3.5.2 Transcription regulation	4,063	2,022	3,0203E-08	4,1047E-07
pspA	Phage shock protein A	3.5.2 Transcription regulation	1,849	0,887	0,00592449	0,01753638
PFCIRM13 8_02505	Helix-turn-helix XRE-family like proteins.	3.5.2 Transcription regulation	11,769	3,557	4,248E-17	2,6556E-15
arsR3	Transcriptional regulators (Bacterial regulatory protein, arsR family)	3.5.2 Transcription regulation	9,029	3,174	0,00028321	0,00126462
tetR1	Transcriptional regulator, TetR family	3.5.2 Transcription regulation	3,006	1,588	7,0767E-05	0,00038613
merR2	Transcriptional regulator, MerR	3.5.2 Transcription regulation	38,449	5,265	8,6754E-07	8,3253E-06
hemR	HemR	3.5.2 Transcription regulation	3,469	1,795	0,00848484	0,02380106
regX	#N/D	3.5.2 Transcription regulation	1,621	0,697	0,02065629	0,04955698
padR	PadR, Transcriptional regulator PadR-like family	3.5.2 Transcription regulation	7,325	2,873	5,5452E-21	8,0885E-19
PFCIRM13 8_06840	transcriptional regulator ArsR	3.5.2 Transcription regulation	6,121	2,614	2,917E-09	4,763E-08
PFCIRM13 8_07310	Transcriptional regulator PadR family	3.5.2 Transcription regulation	3,263	1,706	5,5842E-05	0,00031248
lacl2	Transcriptional regulator, Lacl family protein	3.5.2 Transcription regulation	3,31	1,727	0,00174572	0,00611924
PFCIRM13 8_08435	Two-component system, regulatory protein	3.5.2 Transcription regulation	2,948	1,56	0,00297648	0,00959137



lysR1	Transcriptional regulator, LysR family protein	3.5.2 Transcription regulation	4,075	2,027	7,0714E-06	5,4288E-05
PFCIRM13_8_09520	Regulatory protein, MarR	3.5.2 Transcription regulation	3,825	1,936	0,00578025	0,01720705
PFCIRM13_8_10090	two-component system response regulator	3.5.2 Transcription regulation	2,183	1,126	0,01610133	0,04082237
PFCIRM13_8_10370	anti-sigma factor	3.5.2 Transcription regulation	1,776	0,829	0,00445461	0,01366998
whiB2	transcription factor WhiB	3.5.2 Transcription regulation	2,029	1,021	0,00235321	0,00790909
tetR2	Transcriptional regulator, TetR/AcrR family	3.5.2 Transcription regulation	1,757	0,813	0,01694616	0,04247217
PFCIRM13_8_11280	Silent information regulator protein Sir2 NAD-dependent deacetylase	3.5.2 Transcription regulation	2,118	1,083	0,00929407	0,02561138
PFCIRM13_8_12035	MarR-family transcriptional regulator	3.5.2 Transcription regulation	4,754	2,249	0,00176852	0,00617148
hrcA	Heat-inducible transcription repressor hrcA	3.5.2 Transcription regulation	4,506	2,172	4,2318E-12	1,2183E-10
rpoZ	DNA-directed RNA polymerase omega chain (RNAP omega subunit) (Transcriptase omega chain) (RNA polymerase omega subunit)	3.5.3 Transcription elongation	1,805	0,852	0,00879443	0,02441906
PFCIRM13_8_10430	Transcription termination factor Rho	3.5.4 Transcription termination	2,007	1,005	0,0006899	0,00271004
PFCIRM13_8_10140	DEAD/DEAH box helicase domain protein	3.6 RNA modification	1,599	0,677	0,01455087	0,03763167
trmU	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase	3.6 RNA modification	2,571	1,362	0,00022788	0,00104969
infB	Translation initiation factor IF-2	3.7 Protein synthesis	2,009	1,007	0,00195859	0,00671962
rplS	50S ribosomal protein L19	3.7.1 Ribosomal proteins	1,882	0,912	0,00473468	0,01438816
rpmB	Ribosomal protein L28	3.7.1 Ribosomal proteins	3,235	1,694	2,0683E-05	0,00013389
rpsO	30S ribosomal protein S15	3.7.1 Ribosomal proteins	2,799	1,485	9,3926E-06	6,9429E-05
rplK	50S ribosomal protein L11	3.7.1 Ribosomal proteins	1,802	0,85	0,01296449	0,03397163
rpmA	50S ribosomal protein L27	3.7.1 Ribosomal	2,126	1,088	0,00050893	0,00208527

		proteins				
PFCIRM13_8_04330	translation initiation inhibitor involved in amino acid or purine metabolism	3.7.3 Translation initiation	5,469	2,451	0,00589094	0,01748898
prfA	Peptide chain release factor 1 (RF-1)	3.7.5 Translation termination	2,113	1,079	0,00119823	0,00437685
PFCIRM13_8_06630	Isochorismatase hydrolase	3.7.6 Nonribosomal protein synthesis	2,021	1,015	0,01212049	0,03206726
trxA4	Thioredoxine	3.8 Protein modification	3,519	1,815	2,0155E-05	0,00013125
gst	Glutathione S-transferase	3.8 Protein modification	2,489	1,315	3,3022E-05	0,00019634
trxB	Thioredoxin reductase	3.8 Protein modification	2,696	1,431	0,00071228	0,00278797
PFCIRM13_8_06730	AarF	3.8 Protein modification	3,156	1,658	1,2202E-09	2,1583E-08
PFCIRM13_8_07240	Thioredoxin	3.8 Protein modification	16,925	4,081	5,1511E-05	0,00029198
htrA4	Serine protease	3.8 Protein modification	3,344	1,742	7,2864E-09	1,0995E-07
mrsA	peptide-methionine (S)-S-oxide reductase	3.8 Protein modification	1,818	0,862	0,01140032	0,03048767
PFCIRM13_8_12360	Amino acid-binding ACT	3.8 Protein modification	1,909	0,933	0,00133699	0,00482729
copZ	Copper chaperone CopZ	3.9 Protein folding	2,555	1,353	2,856E-05	0,00017553
hsp20_3	Heat shock protein 20_3 (20 kDa chaperone 3)	3.9 Protein folding	6,775	2,76	2,3131E-15	1,1247E-13
hsp20_1	Heat shock protein 20_1 (20 kDa chaperone 1)	3.9 Protein folding	6,955	2,798	6,2693E-15	2,8578E-13
hsp20_2	Heat shock protein 20_2 (20 kDa chaperone 2)	3.9 Protein folding	2,555	1,353	0,00099365	0,00371643
cspB	Cold shock protein CspB	4.1 Adaptation to atypical conditions	2,664	1,414	2,4027E-06	2,0779E-05
PFCIRM13_8_04080	drug exporters of the RND superfamily	4.1 Adaptation to atypical conditions	1,768	0,822	0,00882506	0,02447305
osmC	OsmC protein	4.1 Adaptation to atypical conditions	2,669	1,416	4,9954E-05	0,00028463
PFCIRM13_8_04225	Multidrug resistance transporter, MFS superfamily protein	4.1 Adaptation to atypical conditions	16,316	4,028	1,252E-05	8,78E-05
PFCIRM13_8_04645	Cyclase family protein	4.1 Adaptation to atypical conditions	35,816	5,163	6,7683E-11	1,5589E-09
katA	Catalase	4.1 Adaptation to	2,129	1,09	0,00057545	0,00232305

		atypical conditions				
dps	Starvation-inducible DNA-binding protein	4.1 Adaptation to atypical conditions	13,789	3,785	7,7027E-21	9,9139E-19
pspC	Possible stress-response transcriptional regulator protein PspC	4.1 Adaptation to atypical conditions	4,094	2,034	2,6719E-08	3,6767E-07
copA	Copper-Cadmimum efflux ATPase	4.2 Detoxification	1,695	0,762	0,00195938	0,00671962
drrA	Daunorubicin resistance ATP-binding protein DrrA	4.2 Detoxification	20,788	4,378	4,2764E-06	3,4913E-05
PFCIRM13 8_03890	Drug resistance transporter	4.2 Detoxification	6,494	2,699	0,00060844	0,00242933
PFCIRM13 8_04715	Aldo/keto reductases, related to diketogulonate reductase	4.2 Detoxification	3,531	1,82	1,9954E-08	2,8167E-07
ahpF	Alkyl hydroperoxide reductase subunit F	4.2 Detoxification	1,883	0,913	0,01456765	0,03763167
PFCIRM13 8_06410	Beta-lactamase-like	4.2 Detoxification	6,756	2,756	0,00266337	0,00886978
matE	Multi antimicrobial extrusion protein MatE	4.2 Detoxification	8,165	3,029	0,00011774	0,0005991
PFCIRM13 8_06835	Hemerythrin HHE cation binding region	4.2 Detoxification	4,86	2,281	5,3397E-11	1,2429E-09
sodA	Iron/Manganese superoxide dismutase (Superoxide dismutase [Mn/Fe]) (SODM)	4.2 Detoxification	9,425	3,237	1,7633E-14	7,0149E-13
PFCIRM13 8_09705	Drug resistance transporter	4.2 Detoxification	4,418	2,143	0,00087606	0,00332205
pcaC1	4-carboxymuconolactone decarboxylase	4.2 Detoxification	7,442	2,896	0,00347011	0,01090891
PFCIRM13 8_00720	Isochorismatase hydrolase	4.3 Antibiotic production	1,742	0,8	0,00459481	0,01400202
PFCIRM13 8_03830	Primase	4.4 Phage-related function	18,861	4,237	1,6746E-32	9,3895E-30
PFCIRM13 8_03840	Gp42	4.4 Phage-related function	6,283	2,651	5,6222E-08	7,0697E-07
PFCIRM13 8_00080	Transposase	4.5 Transposon and IS	7,242	2,856	5,3402E-05	0,00030115
tnpA	Transposase	4.5 Transposon and IS	2,485	1,314	1,721E-05	0,00011411
PFCIRM13 8_06175	Transposase for IS3514b	4.5 Transposon and IS	15,427	3,947	1,0869E-09	1,9985E-08
tnp	Transposase	4.5 Transposon and IS	1,905	0,93	0,00868827	0,02430937

tnp7109-55	Transposase	4.5 Transposon and IS	3,073	1,62	5,4012E-05	0,0003038
TnpD	Transposase IS3/IS911	4.5 Transposon and IS	6,567	2,715	1,9076E-06	1,7106E-05
PFCIRM13 8_00060	Dehydrogenase	4.6 Miscellaneous	6,654	2,734	0,00075901	0,00293411
PFCIRM13 8_02755	acetyltransferase	4.6 Miscellaneous	12,363	3,628	0,00075805	0,00293411
PFCIRM13 8_03090	LMBE-related protein	4.6 Miscellaneous	1,705	0,77	0,01026047	0,02781214
PFCIRM13 8_05685	Zinc-containing alcohol dehydrogenase superfamily	4.6 Miscellaneous	3,604	1,85	0,00585402	0,01740297
PFCIRM13 8_07990	GCN5-related N-acetyltransferase	4.6 Miscellaneous	1,742	0,801	0,00487411	0,01476462
PFCIRM13 8_08590	Methyltransferase	4.6 Miscellaneous	8,446	3,078	0,00074434	0,00288761
PFCIRM13 8_09515	Acetyltransferase	4.6 Miscellaneous	4,163	2,058	0,01734942	0,04313696
PFCIRM13 8_10260	acetyltransferase	4.6 Miscellaneous	5,637	2,495	8,6977E-15	3,8061E-13
PFCIRM13 8_10305	Glycosyl transferase	4.6 Miscellaneous	14,016	3,809	9,6463E-05	0,00050373
PFCIRM13 8_10670	Methyltransferase	4.6 Miscellaneous	3,148	1,654	0,00083633	0,00319352
PFCIRM13 8_10815	acetyltransferase	4.6 Miscellaneous	6,42	2,682	1,804E-05	0,00011925
PFCIRM13 8_11765	Metal dependent phosphohydrolase	4.6 Miscellaneous	11,77	3,557	1,3418E-06	1,2493E-05
obg	GTPase	4.6 Miscellaneous	1,61	0,687	0,01497509	0,03845716
PFCIRM13 8_12295	Metal-dependent hydrolase	4.6 Miscellaneous	3,582	1,841	0,00239099	0,00802374
PFCIRM13 8_00330	protein of unknown function	5.1 Protein of unknown function similar to other proteins from the same organism	58,871	5,879	4,0881E-06	3,3627E-05
PFCIRM13 8_00620	protein of unknown function	5.1 Protein of unknown function similar to other	81,837	6,355	2,8665E-09	4,7158E-08

		proteins from the same organism				
PFCIRM13 8_00960	Hypothetical membrane protein	5.1 Protein of unknown function similar to other proteins from the same organism	1,732	0,792	0,0169054	0,0424186
PFCIRM13 8_03905	protein of unknown function	5.1 Protein of unknown function similar to other proteins from the same organism	5,637	2,495	0,00121989	0,00444855
PFCIRM13 8_05220	Hypothetical secreted protein	5.1 Protein of unknown function similar to other proteins from the same organism	2,715	1,441	1,6849E-05	0,00011309
PFCIRM13 8_00050	Hypothetical membrane protein	5.2 Protein of unknown function similar to proteins from other organisms	7,687	2,942	0,00031592	0,00139926
PFCIRM13 8_00310	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	2,413	1,271	0,00392799	0,01217343
PFCIRM13 8_00355	protein of unknown function	5.2 Protein of unknown function similar to proteins from other organisms	40,353	5,335	5,1216E-07	5,2115E-06
PFCIRM13 8_00410	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	2,872	1,522	8,5136E-06	6,4013E-05
PFCIRM13 8_00535	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	1,689	0,757	0,01559171	0,039707
PFCIRM13 8_00585	protein of unknown function	5.2 Protein of unknown function similar to	20,592	4,364	1,6995E-05	0,00011371

		proteins from other organisms				
PFCIRM13 8_00590	protein of unknown function	5.2 Protein of unknown function similar to proteins from other organisms	2,617	1,388	2,1856E-05	0,00013942
PFCIRM13 8_00605	protein of unknown function	5.2 Protein of unknown function similar to proteins from other organisms	13,578	3,763	5,7204E-05	0,00031848
PFCIRM13 8_00765	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	2,107	1,075	0,00036578	0,00159427
PFCIRM13 8_01260	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	1,81	0,856	0,00204639	0,00697431
PFCIRM13 8_01840	Hypothetical membrane protein	5.2 Protein of unknown function similar to proteins from other organisms	3,313	1,728	4,9837E-06	3,9797E-05
PFCIRM13 8_01925	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	1,81	0,856	0,00837936	0,02353536
PFCIRM13 8_01975	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	2,351	1,233	9,1072E-05	0,00047786
PFCIRM13 8_02050	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	4,451	2,154	5,8951E-10	1,1216E-08
PFCIRM13 8_03095	Hypothetical protein	5.2 Protein of unknown function similar to	2,491	1,317	0,00038423	0,0016582

		proteins from other organisms				
PFCIRM13 8_03570	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	11,847	3,566	1,365E-20	1,6592E-18
PFCIRM13 8_03720	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	2,399	1,262	0,01883697	0,04588504
PFCIRM13 8_03875	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	3,961	1,986	0,0030585	0,00978362
PFCIRM13 8_04010	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	2,499	1,321	2,7596E-05	0,00017008
PFCIRM13 8_04055	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	4,562	2,19	4,0056E-11	9,6311E-10
PFCIRM13 8_04395	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	6,539	2,709	0,00016438	0,00078531
PFCIRM13 8_04790	protein of unknown function	5.2 Protein of unknown function similar to proteins from other organisms	6,498	2,7	1,5613E-15	7,9444E-14
PFCIRM13 8_05045	BioY family protein. Membrane protein	5.2 Protein of unknown function similar to proteins from other organisms	4,248	2,087	2,1616E-08	3,0317E-07
PFCIRM13 8_05295	Hypothetical protein	5.2 Protein of unknown function similar to	2,134	1,094	0,00083032	0,00318167

		proteins from other organisms				
PFCIRM13 8_06350	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	1,977	0,983	0,00401692	0,01243144
PFCIRM13 8_06370	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	1,736	0,796	0,0186894	0,04563884
PFCIRM13 8_06970	protein of unknown function	5.2 Protein of unknown function similar to proteins from other organisms	4,74	2,245	0,00620485	0,01822309
PFCIRM13 8_07000	protein of unknown function	5.2 Protein of unknown function similar to proteins from other organisms	2,702	1,434	7,4494E-05	0,00040245
PFCIRM13 8_07025	protein of unknown function	5.2 Protein of unknown function similar to proteins from other organisms	3,428	1,777	9,4008E-09	1,3993E-07
PFCIRM13 8_07065	protein of unknown function	5.2 Protein of unknown function similar to proteins from other organisms	2,644	1,403	0,00011444	0,00058776
PFCIRM13 8_07110	protein of unknown function	5.2 Protein of unknown function similar to proteins from other organisms	2,558	1,355	0,0001788	0,00084337
PFCIRM13 8_07195	protein of unknown function	5.2 Protein of unknown function similar to proteins from other organisms	15,105	3,917	1,1091E-05	7,9608E-05
PFCIRM13 8_07680	Hypothetical membrane protein	5.2 Protein of unknown function similar to	1,965	0,975	0,00015077	0,00073472



		proteins from other organisms				
PFCIRM13 8_07755	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	1,985	0,989	0,00033375	0,00147225
PFCIRM13 8_08135	cobalt ABC transporter	5.2 Protein of unknown function similar to proteins from other organisms	4,859	2,281	0,01622152	0,04103201
PFCIRM13 8_08205	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	1,835	0,876	0,00220236	0,00748255
PFCIRM13 8_08265	Hypothetical membrane protein	5.2 Protein of unknown function similar to proteins from other organisms	4,718	2,238	0,00757028	0,02162374
PFCIRM13 8_08425	Hypothetical membrane protein	5.2 Protein of unknown function similar to proteins from other organisms	5,262	2,396	0,00296868	0,00958036
PFCIRM13 8_08510	protein of unknown function	5.2 Protein of unknown function similar to proteins from other organisms	2,793	1,482	0,00022247	0,00102691
PFCIRM13 8_08615	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	4,077	2,028	2,6115E-08	3,6164E-07
PFCIRM13 8_08655	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	25,004	4,644	7,6362E-06	5,7813E-05
PFCIRM13 8_08665	Hypothetical membrane protein	5.2 Protein of unknown function similar to	11,764	3,556	0,0001404	0,00069187

		proteins from other organisms				
PFCIRM13 8_08685	Hypothetical membrane protein	5.2 Protein of unknown function similar to proteins from other organisms	17,315	4,114	8,5125E-05	0,00045193
PFCIRM13 8_08690	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	6,558	2,713	0,00303668	0,00972803
PFCIRM13 8_09090	Hypothetical membrane protein	5.2 Protein of unknown function similar to proteins from other organisms	2,448	1,292	2,116E-05	0,00013617
PFCIRM13 8_09425	Hypothetical membrane protein	5.2 Protein of unknown function similar to proteins from other organisms	9,849	3,3	0,0001539	0,0007483
PFCIRM13 8_09600	protein of unknown function	5.2 Protein of unknown function similar to proteins from other organisms	6,055	2,598	2,8939E-17	1,8623E-15
PFCIRM13 8_09605	protein of unknown function	5.2 Protein of unknown function similar to proteins from other organisms	2,469	1,304	2,0402E-05	0,00013246
PFCIRM13 8_10080	Hypothetical transmembrane protein	5.2 Protein of unknown function similar to proteins from other organisms	1,606	0,684	0,00811851	0,0229204
PFCIRM13 8_10150	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	1,779	0,831	0,00957208	0,02624524
PFCIRM13 8_10325	Hypothetical protein	5.2 Protein of unknown function similar to	33,738	5,076	3,0234E-05	0,00018325

		proteins from other organisms				
PFCIRM13 8_10340	protein of unknown function	5.2 Protein of unknown function similar to proteins from other organisms	8,002	3	0,0002369	0,0010776
dedA	protein of unknown function	5.2 Protein of unknown function similar to proteins from other organisms	2,589	1,373	0,0190562	0,04627633
PFCIRM13 8_10635	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	6,25	2,644	1,5135E-14	6,1324E-13
PFCIRM13 8_10675	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	3,512	1,812	0,00010459	0,00054226
PFCIRM13 8_10820	Hypothetical membrane protein	5.2 Protein of unknown function similar to proteins from other organisms	19,805	4,308	2,7134E-05	0,00016771
PFCIRM13 8_11125	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	2,24	1,163	2,328E-05	0,00014637
PFCIRM13 8_11265	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	1,811	0,857	0,00369512	0,01151699
PFCIRM13 8_11285	Hypothetical membrane protein	5.2 Protein of unknown function similar to proteins from other organisms	5,125	2,357	0,00362367	0,01131041
PFCIRM13 8_11965	Hypothetical protein	5.2 Protein of unknown function similar to	6,126	2,615	0,00077752	0,00298983

		proteins from other organisms				
PFCIRM13 8_12135	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	30,564	4,934	1,7357E-06	1,5693E-05
PFCIRM13 8_12190	membrane protein	5.2 Protein of unknown function similar to proteins from other organisms	1,905	0,93	0,00135085	0,00485328
PFCIRM13 8_00045	Hypothetical secreted protein	6 Protein of unknown function, without similarity to other proteins	7,742	2,953	0,00137045	0,00490759
PFCIRM13 8_00340	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	36,263	5,18	1,8097E-06	1,6295E-05
PFCIRM13 8_00350	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	18,263	4,191	1,9215E-05	0,00012625
PFCIRM13 8_00495	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	2,522	1,335	6,3462E-06	4,9415E-05
PFCIRM13 8_00540	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	3,007	1,588	9,3287E-07	8,8745E-06
PFCIRM13 8_00685	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	6,202	2,633	0,00040607	0,00173531
PFCIRM13 8_00775	Hypothetical protein	6 Protein of unknown function, without	1,796	0,845	0,02003396	0,04827567

		similarity to other proteins				
PFCIRM13 8_00860	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	8,159	3,028	0,00100393	0,00374847
PFCIRM13 8_01055	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	3,515	1,813	9,0062E-06	6,7255E-05
PFCIRM13 8_01100	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	40,584	5,343	6,0332E-07	6,0277E-06
PFCIRM13 8_01105	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	3,755	1,909	1,0305E-05	7,5048E-05
PFCIRM13 8_01110	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	2,533	1,341	0,00034502	0,00151891
PFCIRM13 8_01255	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	2,826	1,499	0,01473584	0,03797647
PFCIRM13 8_01415	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	14,972	3,904	1,2887E-05	8,999E-05
PFCIRM13 8_01860	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	3,019	1,594	0,01039925	0,02812554
PFCIRM13 8_01945	Hypothetical protein	6 Protein of unknown function, without	5,918	2,565	0,0004425	0,0018652

		similarity to other proteins				
PFCIRM13 8_01985	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	22,893	4,517	5,3057E-05	0,00029997
PFCIRM13 8_02570	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	2,069	1,049	0,00084363	0,00321579
PFCIRM13 8_02790	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	7,426	2,893	2,7242E-20	3,1372E-18
PFCIRM13 8_02795	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	6,412	2,681	0,0004137	0,00176018
PFCIRM13 8_02945	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	6,159	2,623	0,00644455	0,01887641
PFCIRM13 8_03785	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	8,007	3,001	0,00023102	0,00105967
PFCIRM13 8_03800	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	9,542	3,254	0,00034867	0,0015319
PFCIRM13 8_03815	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	8,461	3,081	2,3555E-17	1,718E-15
PFCIRM13 8_03820	Hypothetical protein	6 Protein of unknown function, without	46,981	5,554	4,7714E-31	2,088E-28

		similarity to other proteins				
PFCIRM13 8_03825	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	38,799	5,278	2,4245E-21	4,509E-19
PFCIRM13 8_03845	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	3,066	1,616	1,6862E-07	1,973E-06
PFCIRM13 8_03915	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	36,161	5,176	3,5589E-07	3,7437E-06
PFCIRM13 8_03920	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	10	3,322	0,00042381	0,00179363
PFCIRM13 8_04020	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	1,733	0,793	0,00990972	0,02700182
PFCIRM13 8_04075	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	5,993	2,583	0,00011576	0,0005904
PFCIRM13 8_04470	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	3,182	1,67	1,1097E-05	7,9608E-05
PFCIRM13 8_04505	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	7,362	2,88	0,00045722	0,0019128
PFCIRM13 8_04700	Hypothetical protein	6 Protein of unknown function, without	1,861	0,896	0,00303494	0,00972803

		similarity to other proteins				
PFCIRM13 8_04710	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	2,875	1,524	2,1237E-05	0,00013626
PFCIRM13 8_04740	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	15,023	3,909	1,6108E-07	1,8949E-06
PFCIRM13 8_04775	protein of unknown function	6 Protein of unknown function, without similarity to other proteins	3,834	1,939	0,0032166	0,01021468
PFCIRM13 8_05060	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	2,658	1,41	0,00327233	0,01036158
PFCIRM13 8_05135	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	8,584	3,102	0,00029126	0,00129528
PFCIRM13 8_05225	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	11,52	3,526	1,0332E-18	9,5632E-17
PFCIRM13 8_05240	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	2,958	1,565	0,01699199	0,04248969
PFCIRM13 8_05605	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	5,103	2,351	0,00026846	0,0012012
PFCIRM13 8_05680	Hypothetical protein	6 Protein of unknown function, without	3,232	1,693	0,00613998	0,01805681



		similarity to other proteins				
PFCIRM13 8_06165	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	5,277	2,4	6,5848E-09	1,0005E-07
PFCIRM13 8_06170	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	7,138	2,836	0,00014478	0,00070711
PFCIRM13 8_06315	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	18,264	4,191	1,4142E-05	9,6699E-05
PFCIRM13 8_06360	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	2,088	1,062	0,00187852	0,00648298
PFCIRM13 8_06635	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	1,534	0,618	0,01811254	0,04467896
PFCIRM13 8_06780	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	29,47	4,881	5,3936E-06	4,2603E-05
PFCIRM13 8_06785	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	10,352	3,372	0,00023019	0,0010581
PFCIRM13 8_06790	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	26,06	4,704	1,4786E-05	0,00010016
PFCIRM13 8_06815	Hypothetical protein	6 Protein of unknown function, without	2,101	1,071	0,0006017	0,0024112

		similarity to other proteins				
PFCIRM13 8_06820	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	8,02	3,004	1,8565E-16	1,0689E-14
PFCIRM13 8_06875	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	2,551	1,351	0,00132553	0,00479382
PFCIRM13 8_06955	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	5,932	2,568	0,00212844	0,00724264
PFCIRM13 8_06990	protein of unknown function	6 Protein of unknown function, without similarity to other proteins	4,812	2,267	8,3428E-08	9,9748E-07
PFCIRM13 8_07105	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	4,688	2,229	3,5802E-13	1,224E-11
PFCIRM13 8_07155	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	6,541	2,71	0,00073586	0,00286486
PFCIRM13 8_07160	protein of unknown function	6 Protein of unknown function, without similarity to other proteins	18,185	4,185	3,7496E-05	0,00021995
PFCIRM13 8_07190	protein of unknown function	6 Protein of unknown function, without similarity to other proteins	10,206	3,351	0,0001789	0,00084337
PFCIRM13 8_07200	protein of unknown function	6 Protein of unknown function, without	9,477	3,244	0,00016107	0,00077454

		similarity to other proteins				
PFCIRM13 8_07255	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	17,52	4,131	2,5936E-05	0,00016168
PFCIRM13 8_07290	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	14,434	3,851	0,00051761	0,00211294
PFCIRM13 8_07475	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	4,931	2,302	0,0041776	0,01285597
PFCIRM13 8_07580	protein of unknown function	6 Protein of unknown function, without similarity to other proteins	7,856	2,974	3,1621E-08	4,2708E-07
PFCIRM13 8_07625	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	14,139	3,822	1,3418E-05	9,2323E-05
PFCIRM13 8_07895	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	2,517	1,332	0,00099312	0,00371643
PFCIRM13 8_07905	protein of unknown function	6 Protein of unknown function, without similarity to other proteins	8,36	3,064	3,7729E-07	3,9498E-06
PFCIRM13 8_08415	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	17,829	4,156	0,00020056	0,00093766
PFCIRM13 8_08440	Hypothetical protein	6 Protein of unknown function, without	11,537	3,528	4,2507E-06	3,4834E-05

		similarity to other proteins				
PFCIRM13 8_08445	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	35,068	5,132	2,2273E-07	2,4613E-06
PFCIRM13 8_08450	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	5,264	2,396	0,00290053	0,00945808
PFCIRM13 8_08455	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	15,732	3,976	7,589E-05	0,00040698
PFCIRM13 8_08535	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	6,391	2,676	0,00179745	0,00626246
PFCIRM13 8_08575	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	4,394	2,136	1,0378E-14	4,4524E-13
PFCIRM13 8_08580	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	4,685	2,228	2,066E-06	1,8375E-05
PFCIRM13 8_08585	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	4,993	2,32	0,00048621	0,00201103
PFCIRM13 8_08680	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	5,209	2,381	0,00161249	0,00568136
PFCIRM13 8_08975	Hypothetical protein	6 Protein of unknown function, without	2,048	1,034	0,01846948	0,04530406

		similarity to other proteins				
PFCIRM13 8_09120	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	6,319	2,66	2,2515E-13	8,2105E-12
PFCIRM13 8_09125	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	10,349	3,371	1,5412E-17	1,1628E-15
PFCIRM13 8_09230	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	12,115	3,599	2,8213E-06	2,4208E-05
PFCIRM13 8_09235	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	12,44	3,637	1,5483E-12	5,0407E-11
PFCIRM13 8_09340	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	9,116	3,188	0,00049039	0,00202096
PFCIRM13 8_09610	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	20,668	4,369	5,2006E-07	5,2437E-06
PFCIRM13 8_09655	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	7,342	2,876	0,00023686	0,0010776
PFCIRM13 8_09660	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	3,091	1,628	3,761E-05	0,00022003
PFCIRM13 8_09675	Hypothetical protein	6 Protein of unknown function, without	3,674	1,877	0,00295749	0,00958036

		similarity to other proteins				
PFCIRM13 8_09695	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	5,343	2,418	0,00257269	0,00859397
PFCIRM13 8_09700	protein of unknown function	6 Protein of unknown function, without similarity to other proteins	28,614	4,839	4,0353E-08	5,2555E-07
PFCIRM13 8_10335	protein of unknown function	6 Protein of unknown function, without similarity to other proteins	4,065	2,023	0,00876188	0,02439059
PFCIRM13 8_10470	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	2,212	1,145	7,5248E-05	0,00040552
PFCIRM13 8_10730	Hypothetical secreted protein	6 Protein of unknown function, without similarity to other proteins	7,306	2,869	9,032E-05	0,00047505
PFCIRM13 8_10740	Hypothetical secreted protein	6 Protein of unknown function, without similarity to other proteins	4,653	2,218	5,8076E-08	7,2612E-07
PFCIRM13 8_10825	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	2,289	1,195	0,01000081	0,02721613
PFCIRM13 8_10945	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	1,587	0,666	0,01389227	0,03605728
PFCIRM13 8_11080	Hypothetical protein	6 Protein of unknown function, without	5,281	2,401	0,00244981	0,00820856

		similarity to other proteins				
PFCIRM13 8_11150	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	1,785	0,836	0,01560696	0,039707
PFCIRM13 8_11315	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	3,456	1,789	0,00118544	0,00433735
PFCIRM13 8_11410	Hypothetical membrane protein	6 Protein of unknown function, without similarity to other proteins	7,12	2,832	0,00039519	0,00169879
PFCIRM13 8_11585	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	5,681	2,506	0,00231847	0,00781634
PFCIRM13 8_11770	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	3,746	1,905	0,00488555	0,01476462
PFCIRM13 8_11920	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	3,976	1,991	0,01790837	0,04427515
PFCIRM13 8_12160	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	4,339	2,117	0,00289119	0,00945188
PFCIRM13 8_12320	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	2,612	1,385	0,01647353	0,0416199
PFCIRM13 8_07965	protein of unknown function	6.2 Protein of unknown function similar to	2,414	1,271	0,00021458	0,0009968

		proteins from other organisms				
PFCIRM13 8_07250	protein of unknown function	7.2 Protein of unknown function similar to proteins from other organisms	2,252	1,171	0,01141197	0,03048767
PFCIRM13 8_08485	protein of unknown function	8.2 Protein of unknown function similar to proteins from other organisms	2,193	1,133	0,00017041	0,00081055
PFCIRM13 8_07140	protein of unknown function	9.2 Protein of unknown function similar to proteins from other organisms	2,139	1,097	0,01049979	0,0283624



**Table S2:** Repressed genes in stationary phase compared to exponential phase.

Name	Descript	Function	FC	1/FC	log2Fold Change	pvalue	padj
rmlC (cps)	dTDP-4-dehydrorhamnose reductase / dTDP-4-dehydrorhamnose 3,5-epimerase	1 Cell envelope and cellular processes	0,349	2,86532951	-1,52	3,4728E-10	7,0388E-09
glmU	UDP-N-acetylglucosamine pyrophosphorylase	1.1 Cell wall	0,63	1,58730159	-0,667	0,01478679	0,03806294
murC	UDP-N-acetylmuramate--L-alanine ligase (UDP-N-acetylmuramoyl-L-alanine synthetase)	1.1 Cell wall	0,425	2,35294118	-1,234	2,2735E-06	1,9898E-05
murD	UDP-N-acetylmuramoylalanine--D-glutamate ligase (UDP-N- acetylmuramoyl-L-alanyl-D-glutamate synthetase) (D-glutamic acid- adding enzyme)	1.1 Cell wall	0,596	1,67785235	-0,746	0,00654238	0,01908631
mraY	Phospho-N-acetylmuramoyl-pentapeptide-transferase (UDP- MurNAc-pentapeptide phosphotransferase)	1.1 Cell wall	0,291	3,43642612	-1,783	2,9208E-10	6,145E-09
murF	UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase (UDP- MurNAc-pentapeptide synthetase) (D-alanyl-D-alanine-adding enzyme)	1.1 Cell wall	0,507	1,97238659	-0,979	0,00069815	0,00273757
murE	UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase (UDP-N-acetylmuramyl-tripeptide synthetase) (Meso- diaminopimelate-adding enzyme) (UDP-MurNAc-tripeptide synthetase)	1.1 Cell wall	0,527	1,89753321	-0,924	0,0006803	0,00268196
ftsI	Cell division protein FtsI (penicillin-binding protein 2) (Peptidoglycan glycosyltransferase)	1.1 Cell wall	0,515	1,94174757	-0,957	0,00032724	0,00144646
PFCIRM1 38_02375	cell-wall peptidases, NlpC/P60 family secreted protein	1.1 Cell wall	0,192	5,20833333	-2,384	3,529E-18	2,8598E-16
ddlA	D-alanine--D-alanine ligase (D-alanylalanine synthetase)	1.1 Cell wall	0,336	2,97619048	-1,576	2,9725E-10	6,1942E-09
dpm	dolichyl-phosphate beta-D-mannosyltransferase	1.1 Cell wall	0,367	2,72479564	-1,448	2,8712E-05	0,00017597
MurA	UDP-N-acetylglucosamine 1-carboxyvinyltransferase (Enoylpyruvate transferase)	1.1 Cell wall	0,555	1,8018018	-0,849	0,00063216	0,00251028
PFCIRM1 38_03660	Rhamnosyltransferase , Glycosyl transferase family 2	1.1 Cell wall	0,597	1,67504188	-0,744	0,00348714	0,01091539
pbpA	Penicillin-binding protein	1.1 Cell wall	0,318	3,14465409	-1,654	1,8129E-07	2,0987E-06
mrc/ponA	penicillin-binding protein (peptidoglycan	1.1 Cell wall	0,645	1,5503876	-0,632	0,01277471	0,03363546

	glycosyltransferase)						
slpG	Surface layer protein G	1.1 Cell wall	0,355	2,81690141	-1,494	1,0062E-08	1,4876E-07
murB	UDP-N-acetylenolpyruvoylglucosamine reductase (UDP-N- acetylmuramate dehydrogenase)	1.1 Cell wall	0,551	1,81488203	-0,859	0,00189094	0,00650532
slpF	surface layer protein F (S-layer protein F)	1.1 Cell wall	0,447	2,23713647	-1,162	1,3114E-05	9,1088E-05
gtfB	Glycosyltransferase	1.1 Cell wall	0,577	1,73310225	-0,794	0,00456797	0,01393965
tagG2	ABC-type polysaccharide export systems	1.1 Cell wall	0,379	2,63852243	-1,401	3,5163E-06	2,9253E-05
tagG1	ABC-type polysaccharide export systems, permease component	1.1 Cell wall	0,518	1,93050193	-0,948	0,00012993	0,00065053
gtfA	Glycosyltransferase, family 2	1.1 Cell wall	0,344	2,90697674	-1,539	4,5625E-08	5,907E-07
glf, rfb	UDP-galactopyranose mutase	1.1 Cell wall	0,472	2,11864407	-1,083	0,00054009	0,00218836
tagO	Undecaprenyl-phosphate alpha-N-acetylglucosaminyltransferase	1.1 Cell wall	0,539	1,85528757	-0,89	0,00115377	0,00425707
PFCIRM1 38_00755	Hypoxanthine/guanine permease 2.A.40.7.2	1.2 Transport/binding proteins and lipoproteins	0,521	1,9193858	-0,94	0,02000989	0,04827567
PFCIRM1 38_00850	Transporter, major facilitator family 2.A.1.52.3	1.2 Transport/binding proteins and lipoproteins	0,181	5,52486188	-2,469	4,6211E-15	2,1513E-13
PFCIRM1 38_01645	ATP-binding protein ABC transporter	1.2 Transport/binding proteins and lipoproteins	0,366	2,73224044	-1,451	0,00309331	0,00986612
PFCIRM1 38_02940	transporter	1.2 Transport/binding proteins and lipoproteins	0,34	2,94117647	-1,557	5,2582E-09	8,2179E-08
PFCIRM1 38_03110	BP protein of monosaccharide ABC transporter (MOS:undef)	1.2 Transport/binding proteins and lipoproteins	0,325	3,07692308	-1,623	1,2191E-05	8,6046E-05
PFCIRM1	Alpha-ketoglutarate permease 2.A.1.6.2	1.2	0,399	2,50626566	-1,326	0,0029664	0,00958036

38_03760		Transport/binding proteins and lipoproteins					
PFCIRM1 38_04250	ABC transporter permease protein	1.2 Transport/binding proteins and lipoproteins	0,496	2,01612903	-1,011	0,00665541	0,01928747
PFCIRM1 38_04695	Xanthine/uracil permease	1.2 Transport/binding proteins and lipoproteins	0,489	2,04498978	-1,031	0,00875932	0,02439059
PFCIRM1 38_04945	Redox protein export	1.2 Transport/binding proteins and lipoproteins	0,62	1,61290323	-0,69	0,01649198	0,0416199
PFCIRM1 38_05530	ATP-binding protein of ABC transporter	1.2 Transport/binding proteins and lipoproteins	0,462	2,16450216	-1,113	4,8661E-05	0,00027872
PFCIRM1 38_05955	ABC transporter ATP-binding protein	1.2 Transport/binding proteins and lipoproteins	0,499	2,00400802	-1,002	0,00328542	0,010388
PFCIRM1 38_07330	polar amino acid ABC transporter, binding protein component	1.2 Transport/binding proteins and lipoproteins	0,45	2,22222222	-1,153	0,00013505	0,00066855
oppD	ABC2 protein of oligopeptide ABC transporter (OPN:undef)	1.2 Transport/binding proteins and lipoproteins	0,283	3,5335689	-1,82	1,2652E-12	4,1943E-11
oppC	IM protein of oligopeptide ABC transporter (OPN:undef)	1.2 Transport/binding proteins and lipoproteins	0,299	3,34448161	-1,742	4,4537E-12	1,2655E-10
oppB	IM protein of oligopeptide ABC transporter (OPN:undef)	1.2	0,358	2,79329609	-1,482	2,4867E-08	3,4656E-07

		Transport/binding proteins and lipoproteins					
PFCIRM1 38_09060	Permease component of cobalt ABC transporter	1.2 Transport/binding proteins and lipoproteins	0,482	2,0746888	-1,054	0,01199629	0,03181561
PFCIRM1 38_09845	BP protein of metallic cation (Zn <sup>2+</sup> Mn <sup>2+</sup> ) ABC transporter	1.2 Transport/binding proteins and lipoproteins	0,395	2,53164557	-1,341	0,00015826	0,00076273
PFCIRM1 38_10105	Preprotein translocase SecA subunit	1.2 Transport/binding proteins and lipoproteins	0,5	2	-1,001	0,00168797	0,00592823
PFCIRM1 38_11820	ABC2 protein of ABC transporter (ART:REG)	1.2 Transport/binding proteins and lipoproteins	0,567	1,76366843	-0,819	0,00148618	0,00527884
tatC	Sec-independent protein translocase protein TatC homolog	1.2.1 Transport/binding of proteins/peptides	0,466	2,14592275	-1,103	1,5475E-05	0,00010451
htrA2	secreted serine protease, trypsin-like serine proteases	1.2.1 Transport/binding of proteins/peptides	0,414	2,41545894	-1,273	0,00044328	0,0018652
bopA	solute binding protein of the ABC transport system	1.2.1 Transport/binding of proteins/peptides	0,471	2,12314225	-1,087	0,00053912	0,00218836
PFCIRM1 38_11405	ATP-binding component ABC transporter	1.2.1 Transport/binding of proteins/peptides	0,572	1,74825175	-0,805	0,01324727	0,03458939
feoB	Ferrous iron uptake protein B 9.A.8.1.x	1.2.3 Transport/binding of inorganic ions	0,367	2,72479564	-1,445	0,00036855	0,00159997
PFCIRM1 38_05990	TrkA, potassium transporter	1.2.3 Transport/binding of	0,636	1,57232704	-0,653	0,00697931	0,02006666

		inorganic ions					
mntH	Mn <sup>2+</sup> and Fe <sup>2+</sup> transporter	1.2.3 Transport/binding of inorganic ions	0,139	7,1942446	-2,852	6,1578E-12	1,6842E-10
cbiO1	Cobalt import ATP-binding protein CbiO	1.2.3 Transport/binding of inorganic ions	0,32	3,125	-1,644	0,00088694	0,00335749
cbiQ2	Cobalt transport protein CbiQ	1.2.3 Transport/binding of inorganic ions	0,306	3,26797386	-1,707	0,00039985	0,00171335
cbiN	Cobalt transport protein CbiN	1.2.3 Transport/binding of inorganic ions	0,266	3,7593985	-1,912	9,1322E-11	2,0389E-09
cbiM	Cobalt transport protein CbiM	1.2.3 Transport/binding of inorganic ions	0,285	3,50877193	-1,811	2,6279E-09	4,3893E-08
PFCIRM1 38_11535	Cation transport protein	1.2.3 Transport/binding of inorganic ions	0,48	2,08333333	-1,06	0,00686414	0,01981363
glpF	Glycerol uptake facilitator protein,	1.2.4 Transport/binding of carbohydrates	0,12	8,33333333	-3,056	3,9947E-12	1,1811E-10
galP	Sodium:galactoside symporter	1.2.4 Transport/binding of carbohydrates	0,403	2,48138958	-1,311	0,00042013	0,0017815
PFCIRM1 38_02775	sugar transporter, major facilitator superfamily	1.2.4 Transport/binding of carbohydrates	0,286	3,4965035	-1,806	0,00280826	0,00922833
PFCIRM1 38_03105	IM protein of monosaccharide ABC transporter (MOS:undef)	1.2.4 Transport/binding of carbohydrates	0,349	2,86532951	-1,52	0,00189062	0,00650532
lldP	L-lactate permease	1.2.4 Transport/binding of carbohydrates	0,257	3,89105058	-1,959	8,415E-10	1,5737E-08
csbX	Alpha-ketoglutarate permease	1.2.4 Transport/binding of	0,309	3,23624595	-1,696	1,2403E-09	2,171E-08

		carbohydrates					
slgT	Sodium/glucose cotransporter (Na+)/glucose symporter) 2.A.21.3.2	1.2.4 Transport/binding of carbohydrates	0,247	4,048583	-2,016	2,1526E-05	0,00013771
PFCIRM1 38_09590	PTS system, Lactose/Cellobiose or galactitol specific IIB subunit	1.2.4 Transport/binding of carbohydrates	0,475	2,10526316	-1,075	0,01727657	0,04300471
gluA	ATP-binding protein of glutamate ABC transporter (PAO : undef : Glutamate )	1.2.5 Transport/binding of amino-acids	0,603	1,65837479	-0,729	0,01697461	0,04248969
PFCIRM1 38_05630	Oligopeptide transporter, OPT family protein.	1.2.5 Transport/binding of amino-acids	0,529	1,89035917	-0,918	0,00026243	0,00118634
PFCIRM1 38_05890	Amino acid permease. membrane protein	1.2.5 Transport/binding of amino-acids	0,425	2,35294118	-1,235	1,715E-05	0,00011405
PFCIRM1 38_06700	ABC-type transport systems, periplasmic component	1.2.5 Transport/binding of amino-acids	0,189	5,29100529	-2,401	2,252E-06	1,9791E-05
metN	Methionine import ATP-binding protein metN	1.2.5 Transport/binding of amino-acids	0,189	5,29100529	-2,405	1,6269E-11	4,0915E-10
PFCIRM1 38_06710	ABC transporter, permease protein	1.2.5 Transport/binding of amino-acids	0,19	5,26315789	-2,397	2,9778E-14	1,143E-12
PFCIRM1 38_07320	polar amino acid ABC transporter, inner membrane component	1.2.5 Transport/binding of amino-acids	0,289	3,46020761	-1,79	3,7466E-12	1,123E-10
PFCIRM1 38_07325	polar amino acid ABC transporter, binding protein	1.2.5 Transport/binding of amino-acids	0,482	2,0746888	-1,053	1,1802E-05	8,3568E-05
PFCIRM1 38_10275	binding protein of oligopeptide ABC transporter (OPN : undef : Oligopeptides)	1.2.5 Transport/binding of amino-acids	0,341	2,93255132	-1,55	8,149E-08	9,7967E-07
PFCIRM1 38_10280	permease protein of oligopeptide ABC transporter	1.2.5 Transport/binding of	0,538	1,85873606	-0,895	0,01001421	0,02721876

		amino-acids					
argK	Kinase ArgK	1.2.5 Transport/binding of amino-acids	0,36	2,77777778	-1,473	6,4451E-05	0,00035521
carA	Carbamoyl-phosphate synthase small chain	1.2.6 Transport/binding of nucleosides, nucleotides, purines and pyrimidines	0,605	1,65289256	-0,726	0,00517053	0,01554
codB	Permease for cytosine/purines, uracil, thiamine, allantoin	1.2.6 Transport/binding of nucleosides, nucleotides, purines and pyrimidines	0,365	2,73972603	-1,452	0,00053094	0,00216329
adk	Adenylate kinase (EC 2.7.4.3) (ATP-AMP transphosphorylase)	1.2.6 Transport/binding of nucleosides, nucleotides, purines and pyrimidines	0,358	2,79329609	-1,482	8,9583E-06	6,7126E-05
PFCIRM1 38_03945	Forkhead-associated protein	1.3 Sensors (signal transduction)	0,634	1,57728707	-0,658	0,01006328	0,02731819
PFCIRM1 38_03950	Forkhead-associated protein	1.3 Sensors (signal transduction)	0,375	2,66666667	-1,416	4,6256E-05	0,00026704
PFCIRM1 38_06685	Sensor protein, ATP-binding region, ATPase-like:Histidine kinase	1.3 Sensors (signal transduction)	0,492	2,03252033	-1,023	0,00116254	0,00428222
PFCIRM1 38_09785	Sensor protein, ATPase-like:Histidine kinase	1.3 Sensors (signal transduction)	0,321	3,1152648	-1,638	1,1946E-10	2,6138E-09
PFCIRM1 38_09790	two-component response regulator	1.3 Sensors (signal transduction)	0,421	2,37529691	-1,249	4,3636E-06	3,5493E-05
sdhB3	Succinate dehydrogenase	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,172	5,81395349	-2,539	9,6739E-11	2,138E-09
sdhA3	Succinate dehydrogenase flavoprotein subunit	1.4 Membrane bioenergetics	0,218	4,58715596	-2,195	2,5141E-11	6,1121E-10

		(electron transport chain and ATP synthase)					
sdhC2	Succinate dehydrogenase cytochrome B-558 subunit	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,227	4,40528634	-2,136	2,2604E-12	6,966E-11
mgtE	magnesium (Mg <sup>2+</sup> ) transporter	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,615	1,62601626	-0,702	0,01194171	0,0317093
fdxA	Ferredoxin	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,35	2,85714286	-1,516	7,4299E-08	8,9815E-07
nuoN	NADH dehydrogenase I chain N	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,165	6,06060606	-2,595	2,7031E-13	9,3881E-12
nuoM	NADH dehydrogenase I chain M	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,157	6,36942675	-2,67	1,707E-12	5,3858E-11
nuoL	NADH dehydrogenase (quinone)	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,201	4,97512438	-2,314	2,59E-09	4,3592E-08
nuoK	NADH dehydrogenase I chain K	1.4 Membrane bioenergetics (electron transport	0,199	5,02512563	-2,329	1,5742E-11	4,005E-10



		chain and ATP synthase)					
nuoJ	NADH-quinone oxidoreductase chain J (NADH dehydrogenase I, chain J)	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,237	4,21940928	-2,075	3,6112E-08	4,7886E-07
nuoI	NADH-quinone oxidoreductase subunit I (NADH dehydrogenase I subunit I) (NDH-1 subunit I)	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,246	4,06504065	-2,023	1,2231E-09	2,1583E-08
nuoH	NADH-quinone oxidoreductase subunit H (NADH dehydrogenase I subunit H)	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,314	3,18471338	-1,673	1,2256E-06	1,1509E-05
nuoG	NADH-quinone oxidoreductase chain G (NADH dehydrogenase I, chain G)	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,288	3,47222222	-1,795	5,0931E-06	4,0523E-05
nuoF	NADH-quinone oxidoreductase chain F (NADH dehydrogenase I, chain F) (NDH-1, chain F)	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,196	5,10204082	-2,354	1,2923E-08	1,8726E-07
nuoE	NADH-quinone oxidoreductase chain E	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,217	4,60829493	-2,206	2,5645E-12	7,7934E-11
nuoD	NADH-quinone oxidoreductase chain D (EC 1.6.99.5) (NADH dehydrogenase I, chain D)	1.4 Membrane bioenergetics (electron transport chain and ATP	0,164	6,09756098	-2,607	4,9573E-12	1,3906E-10

		synthase)					
nuoC	NADH-quinone oxidoreductase chain C (NADH dehydrogenase I, chain C)	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,171	5,84795322	-2,545	6,6857E-15	2,9854E-13
nuoB	NADH-quinone oxidoreductase chain B	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,264	3,78787879	-1,922	5,9342E-12	1,6436E-10
nuoA	NADH-quinone oxidoreductase chain	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,206	4,85436893	-2,276	3,8733E-15	1,8423E-13
PFCIRM1 38_05515	electron transfer oxidoreductase	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,332	3,01204819	-1,589	3,2108E-08	4,31E-07
dhbC	Menaquinone-specific isochorismate synthase	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,549	1,82149362	-0,865	0,00276458	0,00912165
PFCIRM1 38_05930	iron-sulfur protein	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,168	5,95238095	-2,571	2,6088E-13	9,2066E-12
Arth_4141	FAD linked oxidase domain protein	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,284	3,52112676	-1,815	5,2972E-07	5,3166E-06

sdhC1	Succinate dehydrogenase subunit C	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,116	8,62068966	-3,107	4,269E-21	6,6719E-19
sdhA	Succinate dehydrogenase, subunit A	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,113	8,84955752	-3,147	1,8194E-18	1,5923E-16
sdhB	Succinate dehydrogenase, subunit B	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,115	8,69565217	-3,114	5,1373E-16	2,8101E-14
moxR	ATPases MoxR family	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,477	2,09643606	-1,069	0,00012147	0,00061238
atpF	ATP synthase B chain (F0F1 ATP synthase subunit B)	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,554	1,80505415	-0,852	0,00200057	0,00682877
atpH	ATP synthase delta chain	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,509	1,96463654	-0,975	0,0001822	0,00085548
atpA	ATP synthase subunit alpha (ATPase subunit alpha) (ATP synthase F1 sector subunit alpha)	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,412	2,42718447	-1,279	2,3908E-05	0,00014989
atpG	ATP synthase gamma chain (ATP synthase F1 sector	1.4 Membrane	0,41	2,43902439	-1,287	1,1699E-05	8,311E-05

	gamma subunit)	bioenergetics (electron transport chain and ATP synthase)					
atpD	ATP synthase subunit beta (ATPase subunit beta) (ATP synthase F1 sector subunit beta)	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,43	2,3255814	-1,219	0,00046389	0,00193701
atpC	ATP synthase epsilon chain (ATP synthase F1 sector epsilon subunit)	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,477	2,09643606	-1,068	7,1146E-05	0,00038627
nifJ1	Pyruvate synthase/Pyruvate-flavodoxin oxidoreductase	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,271	3,6900369	-1,885	6,8721E-07	6,6828E-06
PFCIRM1 38_10760	FAD-dependent pyridine nucleotide-disulphide oxidoreductase:4Fe-4S ferredoxin, iron-sulfur binding:Aromatic-ring hydroxylase	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,293	3,41296928	-1,77	4,7968E-08	6,1737E-07
PFCIRM1 38_12020	Fe-S oxidoreductases family 2	1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	0,588	1,70068027	-0,766	0,01863881	0,04558667
tatB	Sec-independent protein translocase protein TatB precursor	1.6 Protein secretion	0,489	2,04498978	-1,031	0,0038112	0,01184502
yajC	Preprotein translocase subunit YajC	1.6 Protein secretion	0,608	1,64473684	-0,717	0,0077198	0,02193626
secD	Protein-export membrane protein secD	1.6 Protein secretion	0,397	2,51889169	-1,333	4,5803E-07	4,705E-06
secF	Protein-export membrane protein secF	1.6 Protein secretion	0,196	5,10204082	-2,35	4,9262E-17	2,994E-15
secY	Preprotein translocase SecY subunit	1.6 Protein secretion	0,413	2,42130751	-1,275	2,778E-07	2,965E-06

ftsQ	Cell division protein FtsQ	1.7 Cell division	0,375	2,66666667	-1,415	0,00013391	0,00066741
PFCIRM1 38_04995	ATPase involved in cell division	1.7 Cell division	0,479	2,08768267	-1,061	2,4266E-05	0,0001517
gab	Aldehyde dehydrogenase (Succinate-semialdehyde dehydrogenase) (NAD-dependent aldehyde dehydrogenase)	2.1 Metabolism of carbohydrates and related molecules	0,529	1,89035917	-0,919	0,00593897	0,01753638
pfp (pfk)	pyrophosphate phosphofructokinase	2.1 Metabolism of carbohydrates and related molecules	0,318	3,14465409	-1,652	2,173E-07	2,4253E-06
lpd	Dihydrolipoyl dehydrogenase (E3 component of alpha keto acid dehydrogenase complexes) (Dihydrolipoamide dehydrogenase)	2.1 Metabolism of carbohydrates and related molecules	0,504	1,98412698	-0,989	0,0008952	0,00338291
PFCIRM1 38_04560	Aldose 1-epimerase	2.1 Metabolism of carbohydrates and related molecules	0,357	2,80112045	-1,488	6,0157E-08	7,4363E-07
iolT3	iolT3 (myo-inositol transporter iolT3)	2.1 Metabolism of carbohydrates and related molecules	0,564	1,77304965	-0,827	0,01230528	0,03247763
pta	Phosphate acetyltransferase	2.1 Metabolism of carbohydrates and related molecules	0,23	4,34782609	-2,122	8,9026E-14	3,3015E-12
ackA	Acetate kinase	2.1 Metabolism of carbohydrates and related molecules	0,204	4,90196078	-2,291	8,1723E-19	8,1277E-17
glmS	Glucosamine--fructose-6-phosphate aminotransferase ( Hexosephosphate aminotransferase, D-fructose-6-phosphate amidotransferase)	2.1 Metabolism of carbohydrates and related molecules	0,613	1,63132137	-0,706	0,0068971	0,01988255
glpK	Glycerol kinase (ATP:glycerol 3-phosphotransferase) (Glycerokinase) (GK)	2.1.1 Specific carbohydrate metabolic pathway	0,157	6,36942675	-2,674	4,4048E-10	8,6826E-09
citE	Citrate lyase beta chain (Citrase beta chain) (Citrate(Pro-3S)-lyase beta chain)	2.1.1 Specific carbohydrate metabolic pathway	0,582	1,71821306	-0,782	0,0029122	0,00948198
lacZ	Beta-galactosidase (Lactase) LacZ	2.1.1 Specific carbohydrate metabolic pathway	0,308	3,24675325	-1,699	9,7588E-07	9,2434E-06

glpB	Anaerobic glycerol-3-phosphate dehydrogenase subunit B	2.1.1 Specific carbohydrate metabolic pathway	0,501	1,99600798	-0,998	0,00012019	0,00060873
glpC	Anaerobic glycerol-3-phosphate dehydrogenase subunit C	2.1.1 Specific carbohydrate metabolic pathway	0,462	2,16450216	-1,116	3,0035E-05	0,00018255
otsA	Trehalose-6-phosphate synthase (Alpha,alpha-trehalose-phosphate synthase)	2.1.1 Specific carbohydrate metabolic pathway	0,616	1,62337662	-0,698	0,00748217	0,02142799
rmlA (rfbA)	Glucose-1-phosphate thymidyltransferase	2.1.1 Specific carbohydrate metabolic pathway	0,629	1,58982512	-0,669	0,00798454	0,02261805
galk2	Galactokinase	2.1.1 Specific carbohydrate metabolic pathway	0,541	1,84842884	-0,886	0,00017825	0,00084337
araM	L-arabinose utilization protein	2.1.1 Specific carbohydrate metabolic pathway	0,355	2,81690141	-1,493	8,2974E-05	0,00044172
iolA (msdA)	iolA (Myo-inositol catabolism lolA protein) (Methylmalonic acid semialdehyde dehydrogenase)	2.1.1 Specific carbohydrate metabolic pathway	0,539	1,85528757	-0,893	0,00264533	0,00882316
iolB	iolB (Myo-inositol catabolism lolB protein)	2.1.1 Specific carbohydrate metabolic pathway	0,464	2,15517241	-1,108	0,00048368	0,00200815
PFCIRM1 38_05815	Methylmalonyl-CoA carboxytransferase 5S subunit. (transcarboxylase 5S) 505 bp	2.1.1 Specific carbohydrate metabolic pathway	0,546	1,83150183	-0,873	0,0054605	0,01634416
mmdA	Methylmalonyl-CoA carboxytransferase 12S subunit (EC2.1.3.1) (Transcarboxylase 12S subunit). 610 bp	2.1.1 Specific carbohydrate metabolic pathway	0,417	2,39808153	-1,261	0,00012662	0,00063544
PFCIRM1 38_05825	Methylmalonyl-CoA carboxyltransferase 12S subunit	2.1.1 Specific carbohydrate metabolic pathway	0,447	2,23713647	-1,161	0,00023995	0,00108923
bccp	Methylmalonyl-CoA carboxytransferase , 1.3S subunit (Transcarboxylase, 1.3S subunit). 123bp	2.1.1 Specific carbohydrate metabolic pathway	0,344	2,90697674	-1,54	6,6959E-06	5,1769E-05

galE2	UDP-glucose 4-epimerase	2.1.1 Specific carbohydrate metabolic pathway	0,549	1,82149362	-0,866	0,00038876	0,00167442
PFCIRM1 38_07970	inositol-1-phosphate synthase	2.1.1 Specific carbohydrate metabolic pathway	0,319	3,13479624	-1,65	1,9053E-07	2,1827E-06
treS	Trehalose synthase	2.1.1 Specific carbohydrate metabolic pathway	0,501	1,99600798	-0,997	0,00011277	0,00058058
pccB	Propionyl-CoA carboxylase beta chain	2.1.1 Specific carbohydrate metabolic pathway	0,319	3,13479624	-1,65	6,2064E-10	1,1707E-08
galT	Galactose-1-phosphate uridylyltransferase (UDP-glucose---hexose-1-phosphate uridylyltransferase)	2.1.1 Specific carbohydrate metabolic pathway	0,539	1,85528757	-0,893	0,00560477	0,0167074
tpi1	triosephosphate isomerase 1	2.1.2 Main glycolytic pathways	0,505	1,98019802	-0,985	0,00658728	0,0191917
pgk	Phosphoglycerate kinase	2.1.2 Main glycolytic pathways	0,449	2,22717149	-1,154	0,00023461	0,00107391
pyk1	Pyruvate kinase 1	2.1.2 Main glycolytic pathways	0,47	2,12765957	-1,088	0,00053814	0,00218836
gmp	Phosphoglycerate mutase (D-phosphoglycerate 2,3-phosphomutase)	2.1.2 Main glycolytic pathways	0,507	1,97238659	-0,979	0,00123669	0,0044948
ldh1	L-Lactate dehydrogenase	2.1.2 Main glycolytic pathways	0,466	2,14592275	-1,101	6,1813E-06	4,8476E-05
fba1	Fructose-bisphosphate aldolase class II	2.1.2 Main glycolytic pathways	0,53	1,88679245	-0,915	0,00770202	0,02191421
pfkA	6-phosphofructokinase	2.1.2 Main glycolytic pathways	0,457	2,18818381	-1,13	7,0413E-05	0,00038516
gmpB	phosphoglycerate mutase	2.1.2 Main glycolytic pathways	0,464	2,15517241	-1,107	0,00493771	0,01488114
mdh	Malate dehydrogenase	2.1.3 TCA cycle	0,522	1,91570881	-0,938	0,00276718	0,00912165
fumC	Fumarate hydratase, class-II	2.1.3 TCA cycle	0,47	2,12765957	-1,089	0,00044967	0,00188482
icd	Putative isocitrate/isopropylmalate dehydrogenase	2.1.3 TCA cycle	0,485	2,06185567	-1,045	0,0004114	0,00175468
cat	Coenzyme A transferase (Putative succinyl-CoA or	2.1.4 Substrate-	0,421	2,37529691	-1,247	0,00011893	0,00060373

	butyryl-CoA:coenzyme A transferase)	specific entries to carbohydrate metabolic pathway					
rmIB (rfbB)	DTDP-glucose 4,6-dehydratase	2.1.4 Substrate-specific entries to carbohydrate metabolic pathway	0,492	2,03252033	-1,023	6,695E-05	0,00036806
argH	Argininosuccinate lyase (Arginosuccinase)	2.2 Metabolism of amino acids and related molecules	0,479	2,08768267	-1,061	3,7005E-05	0,00021766
argJ	Arginine biosynthesis bifunctional protein argJ (Glutamate N-acetyltransferase) (Ornithine acetyltransferase) (Ornithine transacetylase) (OATase)	2.2 Metabolism of amino acids and related molecules	0,347	2,88184438	-1,526	2,9771E-07	3,162E-06
argC	N-acetyl-gamma-glutamyl-phosphate reductase (AGPR) (N- acetyl-glutamate semialdehyde dehydrogenase) (NAGSA dehydrogenase)	2.2 Metabolism of amino acids and related molecules	0,256	3,90625	-1,963	1,7231E-12	5,3858E-11
hisG	ATP phosphoribosyltransferase	2.2 Metabolism of amino acids and related molecules	0,633	1,57977883	-0,66	0,00677543	0,01960931
hisE	Phosphoribosyl-ATP pyrophosphohydrolase	2.2 Metabolism of amino acids and related molecules	0,501	1,99600798	-0,998	0,00289431	0,00945188
trpD	Anthranilate phosphoribosyltransferase	2.2 Metabolism of amino acids and related molecules	0,271	3,6900369	-1,882	4,6178E-13	1,5544E-11
PFCIRM1 38_02155	Lysine decarboxylase	2.2 Metabolism of amino acids and related molecules	0,463	2,15982721	-1,112	1,1428E-05	8,1446E-05
PFCIRM1 38_02175	aminotransferase	2.2 Metabolism of amino acids and related molecules	0,317	3,15457413	-1,658	8,9783E-11	2,0252E-09
hisC	Histidinol-phosphate aminotransferase (Imidazole acetol-phosphate transaminase)	2.2 Metabolism of amino acids and related molecules	0,518	1,93050193	-0,948	0,00016639	0,00079316
hisB	Imidazoleglycerol-phosphate dehydratase (IGPD)	2.2 Metabolism of amino acids and	0,44	2,27272727	-1,186	0,00010079	0,00052382



		related molecules					
hisA	Phosphoribosyl isomerase A (1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino]imidazole-4-carboxamide isomerase) (Phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase) (N-(5-phosphoribosyl)anthranilate isomerase) (PRAI)	2.2 Metabolism of amino acids and related molecules	0,5	2	-1,001	0,00037469	0,00162341
carB (pyrA)	Carbamoyl-phosphate synthase large chain (Carbamoyl-phosphate synthetase ammonia chain)	2.2 Metabolism of amino acids and related molecules	0,397	2,51889169	-1,332	5,4695E-05	0,00030685
metH	methionine synthase (5-methyltetrahydrofolate:L-homocysteine S-methyltransferase)	2.2 Metabolism of amino acids and related molecules	0,474	2,10970464	-1,076	1,9411E-05	0,00012716
ilvC	Ketol-acid reductoisomerase (Acetohydroxy-acid isomeroreductase) (Alpha-keto-beta-hydroxylacil reductoisomerase)	2.2 Metabolism of amino acids and related molecules	0,208	4,80769231	-2,262	6,3096E-17	3,7312E-15
ilvD	Dihydroxy-acid dehydratase	2.2 Metabolism of amino acids and related molecules	0,46	2,17391304	-1,122	2,6212E-05	0,00016287
asnB	Asparagine synthase (Glutamine-hydrolyzing)	2.2 Metabolism of amino acids and related molecules	0,672	1,48809524	-0,574	0,0171151	0,04270219
ansA	L-asparaginase I	2.2 Metabolism of amino acids and related molecules	0,472	2,11864407	-1,082	8,953E-05	0,00047203
ask	Aspartokinase (Aspartate kinase)	2.2 Metabolism of amino acids and related molecules	0,557	1,79533214	-0,845	0,00405315	0,01252584
cys2	cysteine synthase 2	2.2 Metabolism of amino acids and related molecules	0,057	17,5438596	-4,135	1,4784E-14	6,1033E-13
PFCIRM1 38_09945	alanine racemase	2.2 Metabolism of amino acids and related molecules	0,444	2,25225225	-1,172	0,00146153	0,00519973
tyrB	Aspartate transaminase (Aminotransferase)	2.2 Metabolism of amino acids and related molecules	0,418	2,3923445	-1,259	3,3558E-05	0,00019898

gabT	4-aminobutyrate aminotransferase	2.2 Metabolism of amino acids and related molecules	0,152	6,57894737	-2,714	5,2032E-19	5,4212E-17
lysA	Diaminopimelate decarboxylase (DAP decarboxylase)	2.2 Metabolism of amino acids and related molecules	0,552	1,8115942	-0,857	0,00051744	0,00211294
thrA/hom	Homoserine dehydrogenase	2.2 Metabolism of amino acids and related molecules	0,411	2,43309002	-1,283	1,7656E-07	2,0549E-06
serB/thrH	Phosphoserine phosphatase/homoserine phosphotransferase bifunctional protein	2.2 Metabolism of amino acids and related molecules	0,36	2,77777778	-1,475	5,9958E-08	7,4363E-07
bkdB	dihydrolipoyllysine-residue (2-methylpropanoyl)transferase. Lipoamide acyltransferase component of branched-chain alpha-keto acid dehydrogenase complex (E2) (Dihydrolipoamide branched chain transacylase)	2.2 Metabolism of amino acids and related molecules	0,375	2,66666667	-1,416	2,0794E-07	2,3399E-06
bkdA2	2-oxoisovalerate dehydrogenase subunit beta (EC 1.2.4.4) (Branched- chain alpha-keto acid dehydrogenase E1 component beta chain) (BCKDH E1-beta) Pyruvate dehydrogenase E1 component subunit beta	2.2 Metabolism of amino acids and related molecules	0,341	2,93255132	-1,554	1,5683E-08	2,2428E-07
bkdA1	2-oxoisovalerate dehydrogenase subunit alpha (Branched- chain alpha-keto acid dehydrogenase E1 component alpha chain) (BCKDH E1-alpha)	2.2 Metabolism of amino acids and related molecules	0,524	1,90839695	-0,932	0,00010967	0,00056593
argG	Argininosuccinate synthase	2.2 Metabolism of amino acids and related molecules	0,324	3,08641975	-1,625	2,3102E-07	2,5274E-06
PFCIRM1 38_11035	Chorismate mutase	2.2 Metabolism of amino acids and related molecules	0,446	2,24215247	-1,166	7,3782E-05	0,00039959
cysE	Serine acetyltransferase	2.2 Metabolism of amino acids and related molecules	0,55	1,81818182	-0,863	0,00324044	0,01027548
cys1	Cysteine synthase 1	2.2 Metabolism of amino acids and related molecules	0,585	1,70940171	-0,774	0,00184305	0,0063706

prs	Ribose-phosphate pyrophosphokinase	2.3 Metabolism of nucleotides and nucleic acids	0,337	2,96735905	-1,571	3,3471E-09	5,3455E-08
pdxT	Glutamine amidotransferase subunit pdxT (Glutamine amidotransferase glutaminase subunit pdxT)	2.3 Metabolism of nucleotides and nucleic acids	0,44	2,27272727	-1,184	0,00300551	0,00967066
purM	Phosphoribosylformylglycinamide cyclo-ligase	2.3 Metabolism of nucleotides and nucleic acids	0,628	1,59235669	-0,67	0,01564276	0,03975188
rnd	Ribonuclease D (3 -5 exonuclease)	2.3 Metabolism of nucleotides and nucleic acids	0,386	2,59067358	-1,374	7,0629E-08	8,629E-07
apt	Adenine phosphoribosyltransferase (AMP:diphosphate phospho-D-ribosyltransferase)	2.3 Metabolism of nucleotides and nucleic acids	0,32	3,125	-1,645	4,6044E-07	4,7077E-06
pyrD	Dihydroorotate dehydrogenase 2	2.3 Metabolism of nucleotides and nucleic acids	0,537	1,86219739	-0,898	0,00077684	0,00298983
upp	Uracil phosphoribosyltransferase (UMP pyrophosphorylase) (UPRTase)	2.3 Metabolism of nucleotides and nucleic acids	0,526	1,90114068	-0,926	0,01783752	0,04414987
PFCIRM1 38_04525	ribose-phosphate pyrophosphokinase (RPPK) (Phosphoribosyl pyrophosphate synthetase) (P-Rib-PP synthetase) (PRPP synthetase)	2.3 Metabolism of nucleotides and nucleic acids	0,411	2,43309002	-1,283	4,5375E-07	4,683E-06
pyrE	Orotate phosphoribosyltransferase (OPRT) (OPRTase)	2.3 Metabolism of nucleotides and nucleic acids	0,602	1,66112957	-0,733	0,01344317	0,03505799
hpt	Hypoxanthine phosphoribosyltransferase	2.3 Metabolism of nucleotides and nucleic acids	0,324	3,08641975	-1,628	3,8731E-08	5,105E-07
gpt	Putative purine phosphoribosyltransferase	2.3 Metabolism of nucleotides and nucleic acids	0,444	2,25225225	-1,17	2,2522E-06	1,9791E-05
PFCIRM1 38_10750	Dihydroorotate dehydrogenase	2.3 Metabolism of nucleotides and nucleic acids	0,272	3,67647059	-1,88	9,3166E-06	6,9101E-05

guaB2	Inosine-5 -monophosphate dehydrogenase (IMP dehydrogenase) (IMPDH) (IMPD) / GMP reductase	2.3 Metabolism of nucleotides and nucleic acids	0,476	2,10084034	-1,07	0,00014279	0,00070049
purH	Phosphoribosylaminoimidazolecarboxamide formyltransferase (AICAR transformylase)	2.3 Metabolism of nucleotides and nucleic acids	0,398	2,51256281	-1,33	3,3022E-05	0,00019634
ugpQ2	Glycerophosphoryl diester phosphodiesterase	2.4 Metabolism of lipids	0,241	4,14937759	-2,052	3,7194E-10	7,4661E-09
cdsA	Phosphatidate cytidyltransferase	2.4 Metabolism of lipids	0,516	1,9379845	-0,955	0,00058516	0,0023579
inhA	Enoyl-[acyl-carrier-protein] reductase (NADH)	2.4 Metabolism of lipids	0,591	1,69204738	-0,758	0,00825452	0,02327434
fabG	3-oxoacyl-[acyl-carrier protein] reductase	2.4 Metabolism of lipids	0,529	1,89035917	-0,92	0,00967845	0,02650153
plsC3	Acyltransferase PlsC	2.4 Metabolism of lipids	0,592	1,68918919	-0,756	0,00480094	0,01456929
PFCIRM1 38_02490	Phosphatidylglycerophosphate synthase	2.4 Metabolism of lipids	0,631	1,58478605	-0,665	0,01122571	0,0301003
pgs2	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyl-transferase	2.4 Metabolism of lipids	0,376	2,65957447	-1,409	0,00013478	0,00066855
PPA2234	Long-chain-fatty-acid--CoA ligase/synthetase	2.4 Metabolism of lipids	0,574	1,74216028	-0,802	0,0029171	0,00948383
pf1861	Carboxylic ester hydrolase	2.4 Metabolism of lipids	0,485	2,06185567	-1,045	9,3149E-05	0,00048759
fabF (fabB)	3-oxoacyl-[acyl-carrier-protein] synthase ( Beta-ketoacyl-ACP synthase)	2.4 Metabolism of lipids	0,159	6,28930818	-2,652	2,6775E-17	1,8307E-15
acpP	Acyl carrier protein (ACP)	2.4 Metabolism of lipids	0,169	5,91715976	-2,569	1,049E-18	9,5632E-17
fabH	3-oxoacyl-(Acyl-carrier-protein) synthase III	2.4 Metabolism of lipids	0,145	6,89655172	-2,782	1,2113E-22	2,9449E-20
fabD	Malonyl CoA-acyl carrier protein transacylase	2.4 Metabolism of lipids	0,271	3,6900369	-1,884	4,0889E-12	1,1929E-10
accA	Acetyl-CoA carboxylase	2.4 Metabolism of lipids	0,598	1,6722408	-0,742	0,00763265	0,0217451
PFCIRM1 38_00065	2-dehydropantoate 2-reductase	2.5 Metabolism of coenzymes and	0,324	3,08641975	-1,624	4,8218E-09	7,645E-08

		prosthetic groups					
pntB	NADH dehydrogenase	2.5 Metabolism of coenzymes and prosthetic groups	0,345	2,89855072	-1,533	4,9616E-06	3,9766E-05
pntA	NAD(P)(+) transhydrogenase (AB-specific).	2.5 Metabolism of coenzymes and prosthetic groups	0,331	3,02114804	-1,595	9,0372E-07	8,6347E-06
coaD	Phosphopantetheine adenylyltransferase (Pantetheine-phosphate adenylyltransferase) (PPAT) (Dephospho-CoA pyrophosphorylase)	2.5 Metabolism of coenzymes and prosthetic groups	0,55	1,81818182	-0,862	0,00465722	0,01417245
dxr	1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXP reductoisomerase) (1-deoxyxylulose-5-phosphate reductoisomerase) (2-C- methyl-D-erythritol 4-phosphate synthase)	2.5 Metabolism of coenzymes and prosthetic groups	0,569	1,75746924	-0,814	0,00346053	0,01089446
pdxS	Pyridoxal biosynthesis lyase pdxS	2.5 Metabolism of coenzymes and prosthetic groups	0,342	2,92397661	-1,549	2,0466E-07	2,3202E-06
thiL	Thiamine monophosphate kinase	2.5 Metabolism of coenzymes and prosthetic groups	0,61	1,63934426	-0,714	0,00273067	0,00905258
cobU	CobU Bifunctional cobalamin biosynthesis pyrophosphate enzyme	2.5 Metabolism of coenzymes and prosthetic groups	0,355	2,81690141	-1,493	0,00937282	0,02575884
cbiA	Cobyrinic acid A,C-diamide synthase, CbiA	2.5 Metabolism of coenzymes and prosthetic groups	0,487	2,05338809	-1,037	0,003179	0,01010995
cobA2	Cob(I)alamin adenosyltransferase	2.5 Metabolism of coenzymes and prosthetic groups	0,414	2,41545894	-1,273	0,01829595	0,04502985
cbiP/cobQ	Cobyrinic acid synthase CbiP/CobQ	2.5 Metabolism of coenzymes and prosthetic groups	0,464	2,15517241	-1,108	0,00590639	0,01751107
idsA	Heptaprenyl diphosphate synthase component II	2.5 Metabolism of coenzymes and prosthetic groups	0,194	5,15463918	-2,363	2,839E-17	1,8623E-15

hemA	Glutamyl-tRNA reductase	2.5 Metabolism of coenzymes and prosthetic groups	0,299	3,34448161	-1,741	5,6016E-09	8,6311E-08
hemC	Porphobilinogen deaminase (PBG) (Hydroxymethylbilane synthase) (HMBS) (Pre-uroporphyrinogen synthase)	2.5 Metabolism of coenzymes and prosthetic groups	0,332	3,01204819	-1,59	0,00125684	0,00456048
hemE	Uroporphyrinogen decarboxylase	2.5 Metabolism of coenzymes and prosthetic groups	0,381	2,62467192	-1,391	0,00492605	0,01486647
hemY	Protoporphyrinogen oxidase (PPO) HemY	2.5 Metabolism of coenzymes and prosthetic groups	0,314	3,18471338	-1,671	0,00044139	0,00186438
hemB	Delta-aminolevulinic acid dehydratase (Porphobilinogen synthase)	2.5 Metabolism of coenzymes and prosthetic groups	0,267	3,74531835	-1,903	4,8843E-05	0,00027903
ribC (ribE)	Riboflavin synthase alpha chain	2.5 Metabolism of coenzymes and prosthetic groups	0,494	2,0242915	-1,017	0,00267583	0,00889775
ribAB	GTP cyclohydrolase II protein, Riboflavin biosynthesis protein	2.5 Metabolism of coenzymes and prosthetic groups	0,402	2,48756219	-1,316	0,00013753	0,00067926
cobA	CobA Uroporphyrinogen III methyltransferase	2.5 Metabolism of coenzymes and prosthetic groups	0,312	3,20512821	-1,678	7,9359E-05	0,00042454
chID	ChID, Mg-chelatase subunit ChID	2.5 Metabolism of coenzymes and prosthetic groups	0,5	2	-1,001	0,01779438	0,04409299
dkgA	2,5-diketo-D-gluconate reductase A	2.5 Metabolism of coenzymes and prosthetic groups	0,447	2,23713647	-1,161	0,00047263	0,00196598
bluB/cobT 2	Phosphoribosyltransferase/nitroreductase (fusion gene) (Nicotinate-nucleotide-dimethylbenzimidazole phosphoribosyltransferase)	2.5 Metabolism of coenzymes and prosthetic groups	0,339	2,94985251	-1,562	1,9317E-08	2,7445E-07
folD	Methylenetetrahydrofolate dehydrogenase (Bifunctional protein)	2.5 Metabolism of coenzymes and prosthetic groups	0,397	2,51889169	-1,333	2,2664E-07	2,4919E-06

menB	Naphthoate synthase (Dihydroxynaphthoic acid synthetase)	2.5 Metabolism of coenzymes and prosthetic groups	0,46	2,17391304	-1,12	2,9891E-05	0,00018218
cbiL	CbiL Precorrin-2 C20-methyltransferase	2.5 Metabolism of coenzymes and prosthetic groups	0,43	2,3255814	-1,218	0,00013492	0,00066855
cbiF	CbiF Precorrin-4 C11-methyltransferase	2.5 Metabolism of coenzymes and prosthetic groups	0,332	3,01204819	-1,592	1,5932E-05	0,00010726
cysG/cbiX	CysG/CbiX	2.5 Metabolism of coenzymes and prosthetic groups	0,441	2,2675737	-1,181	0,00086698	0,00329333
folC	Folypolyglutamate synthase (bifunctional enzyme)	2.5 Metabolism of coenzymes and prosthetic groups	0,656	1,52439024	-0,608	0,01364176	0,03553355
PFCIRM1 38_06220	NUDIX hydrolase	2.6 Metabolism of phosphate	0,392	2,55102041	-1,353	1,5664E-06	1,434E-05
PFCIRM1 38_08215	Inorganic pyrophosphatase	2.6 Metabolism of phosphate	0,478	2,09205021	-1,064	0,00076125	0,0029376
cysW	Sulfate transport system permease protein CysW	2.7 Metabolism of sulfur	0,081	12,345679	-3,62	2,4676E-07	2,6862E-06
cysT	Sulphate transport system permease protein 2:Sulfate ABC transporter, permease protein CysT	2.7 Metabolism of sulfur	0,097	10,3092784	-3,368	6,5968E-07	6,4725E-06
PFCIRM1 38_00925	Sulfate-binding protein precursor	2.7 Metabolism of sulfur	0,118	8,47457627	-3,078	4,7385E-05	0,00027212
nirA2/sir2	Sulfite reductase [ferredoxin]	2.7 Metabolism of sulfur	0,076	13,1578947	-3,717	1,1963E-09	2,1456E-08
cysH	Putative phosphoadenosine phosphosulfate reductase	2.7 Metabolism of sulfur	0,051	19,6078431	-4,3	2,4806E-17	1,7508E-15
mmcV/cys D	Sulfate adenylyltransferase subunit 2 (Sulfate adenylyltransferase) (SAT) (ATP-sulfurylase small subunit) (Mitomycin biosynthesis protein V)	2.7 Metabolism of sulfur	0,035	28,5714286	-4,827	7,1865E-27	2,2463E-24
cysN	Sulfate adenylyltransferase, large subunit/ ATP-sulfurylase, subunit 1 (ATP:sulfate adenylyltransferase)	2.7 Metabolism of sulfur	0,033	30,3030303	-4,932	1,7165E-32	9,3895E-30
norB	Nitric-oxide reductase subunit B (nitric-oxide reductase)	2.8 Metabolism of nitrogen/nitrate and	0,475	2,10526316	-1,073	0,00068892	0,00271004

		nitrite					
glnA1	Glutamine synthetase	2.8 Metabolism of nitrogen/nitrate and nitrite	0,543	1,84162063	-0,882	0,00408826	0,01261651
dnaN	DNA polymerase III, beta chain	3.1 DNA replication	0,493	2,02839757	-1,019	0,0005583	0,00225798
PFCIRM1 38_07260	Exonuclease	3.1 DNA replication	0,393	2,54452926	-1,346	5,3832E-06	4,2603E-05
PFCIRM1 38_09150	DeaD/DeaH box helicase	3.1 DNA replication	0,583	1,71526587	-0,779	0,00197229	0,00675333
pepP	Xaa-Pro aminopeptidase I	3.10 Protein degradation	0,583	1,71526587	-0,78	0,00361727	0,01130657
PFCIRM1 38_02300	Zinc metallopeptidase	3.10 Protein degradation	0,562	1,77935943	-0,831	0,00608512	0,01791958
PFCIRM1 38_03755	Amidohydrolase (Peptidase M20D) (Putative metal-dependent amidase/aminoacylase/carboxypeptidase)	3.10 Protein degradation	0,418	2,3923445	-1,26	0,00551416	0,01648223
map	Methionine aminopeptidase (MAP) (Peptidase M)	3.10 Protein degradation	0,315	3,17460317	-1,669	1,4872E-06	1,3673E-05
PFCIRM1 38_10230	Secreted protease with a PDZ domain	3.10 Protein degradation	0,312	3,20512821	-1,682	1,7693E-11	4,3992E-10
pepN	Aminopeptidase N , Lysyl aminopeptidase	3.10 Protein degradation	0,521	1,9193858	-0,94	0,0039231	0,01217343
clpP2	ATP-dependent Clp protease proteolytic subunit 2 (Endopeptidase Clp 2)	3.10 Protein degradation	0,583	1,71526587	-0,777	0,0181493	0,04471922
PFCIRM1 38_05540	nuclease (RecB family)	3.2 DNA restriction and modification (and repair)	0,639	1,56494523	-0,646	0,01158533	0,030892
nth	Putative endonuclease III	3.2 DNA restriction and modification (and repair)	0,536	1,86567164	-0,9	0,00227571	0,0076968
PFCIRM1 38_06795	Hypothetical protein	3.2 DNA restriction and modification (and repair)	0,466	2,14592275	-1,1	0,00888308	0,02460276
recB	Exodeoxyribonuclease V beta chain RecB	3.3 DNA recombination, and repair	0,657	1,52207002	-0,607	0,01511099	0,03871528



PFCIRM1 38_11040	transcription initiation factor	3.5.1 Transcription initiation	0,631	1,58478605	-0,664	0,00899062	0,02486914
rnc	Ribonuclease III	3.5.2 Transcription regulation	0,615	1,62601626	-0,701	0,00989615	0,02699846
dskA	Transcriptional regulators, TraR/DksA family	3.5.2 Transcription regulation	0,433	2,30946882	-1,206	7,5969E-06	5,7716E-05
PFCIRM1 38_01955	Response regulator receiver protein	3.5.2 Transcription regulation	0,417	2,39808153	-1,262	1,9882E-06	1,7755E-05
merR1	Transcriptional regulator, MerR	3.5.2 Transcription regulation	0,587	1,70357751	-0,767	0,00113456	0,00419328
PFCIRM1 38_02925	Transcriptional regulator, MerR	3.5.2 Transcription regulation	0,479	2,08768267	-1,062	0,00026844	0,0012012
deoR3	Putative regulatory protein, DeoR family	3.5.2 Transcription regulation	0,503	1,98807157	-0,991	0,00029069	0,00129528
PFCIRM1 38_04635	Transcriptional Regulator, GntR family protein	3.5.2 Transcription regulation	0,39	2,56410256	-1,359	5,9982E-06	4,7209E-05
PFCIRM1 38_06690	two component system response regulator	3.5.2 Transcription regulation	0,27	3,7037037	-1,888	5,6001E-10	1,0748E-08
PFCIRM1 38_08370	Two-component system response regulator	3.5.2 Transcription regulation	0,447	2,23713647	-1,162	6,7331E-06	5,1873E-05
sirR	Iron dependent repressor	3.5.2 Transcription regulation	0,232	4,31034483	-2,106	2,1627E-10	4,5941E-09
PFCIRM1 38_09335	Excisionase	3.5.2 Transcription regulation	0,375	2,66666667	-1,415	1,4199E-06	1,3109E-05
PFCIRM1 38_10835	two-component response regulator LuxR	3.5.2 Transcription regulation	0,544	1,83823529	-0,878	0,0085986	0,0240893
PFCIRM1 38_10890	Cell envelope-related transcriptional attenuator	3.5.2 Transcription regulation	0,652	1,53374233	-0,618	0,00875209	0,02439059
rpoC	DNA-directed RNA polymerase beta chain (RNAP beta subunit) (Transcriptase beta chain) (RNA polymerase beta subunit)	3.5.3 Transcription elongation	0,502	1,99203187	-0,994	0,00938291	0,02575884
rpoA	DNA-directed RNA polymerase alpha chain (RNAP alpha subunit) (Transcriptase alpha chain) (RNA polymerase subunit alpha)	3.5.3 Transcription elongation	0,253	3,95256917	-1,983	1,0341E-09	1,9174E-08
nusG	Transcription antitermination protein NusG	3.5.4 Transcription termination	0,57	1,75438596	-0,81	0,00287161	0,00940582

spoU1	rRNA methylase	3.6 RNA modification	0,392	2,55102041	-1,35	6,7802E-05	0,0003718
PFCIRM1 38_05905	tRNA/rRNA methyltransferase	3.6 RNA modification	0,481	2,07900208	-1,057	0,00020819	0,00096917
rsmC	rRNA (guanine-N2-)-methyltransferase	3.6 RNA modification	0,589	1,69779287	-0,764	0,01885313	0,04588504
PFCIRM1 38_10045	GatB_Yqey, GatB/Yqey	3.6 RNA modification	0,651	1,53609831	-0,62	0,01167347	0,0310347
orn	Oligoribonuclease	3.6 RNA modification	0,33	3,03030303	-1,6	4,508E-10	8,8068E-09
gatC	Glutamyl-tRNA(Gln) amidotransferase subunit C (Aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit C)	3.7 Protein synthesis	0,437	2,28832952	-1,194	8,5304E-05	0,00045193
gatB	Aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit B (Asp/Glu-ADT subunit B)	3.7 Protein synthesis	0,343	2,9154519	-1,544	3,9131E-08	5,1268E-07
rplY	Ribosomal protein L25	3.7.1 Ribosomal proteins	0,308	3,24675325	-1,701	1,4762E-09	2,5634E-08
rplT	50S ribosomal protein L20	3.7.1 Ribosomal proteins	0,276	3,62318841	-1,858	5,1494E-08	6,5548E-07
rplM	50S ribosomal protein L35	3.7.1 Ribosomal proteins	0,331	3,02114804	-1,593	2,713E-09	4,497E-08
rpsA	30S ribosomal protein S1	3.7.1 Ribosomal proteins	0,503	1,98807157	-0,992	0,00452294	0,01384083
rpsR	30S ribosomal protein S18	3.7.1 Ribosomal proteins	0,487	2,05338809	-1,038	0,0007434	0,00288761
rplI	50S ribosomal protein L9	3.7.1 Ribosomal proteins	0,46	2,17391304	-1,12	0,00066158	0,00261288
rpsK	30S ribosomal protein S11	3.7.1 Ribosomal proteins	0,493	2,02839757	-1,019	0,00059883	0,00240409
rplA	50S ribosomal protein L1	3.7.1 Ribosomal proteins	0,536	1,86567164	-0,899	0,0063055	0,01849387
rplJ	50S ribosomal protein L10	3.7.1 Ribosomal proteins	0,064	15,625	-3,963	2,5466E-21	4,509E-19
rplL	50S ribosomal protein L7/L12	3.7.1 Ribosomal proteins	0,063	15,8730159	-3,991	2,679E-21	4,509E-19

rpsG	30S ribosomal protein S7	3.7.1 Ribosomal proteins	0,625	1,6	-0,678	0,01500913	0,03849939
rpsJ	30S ribosomal protein S10	3.7.1 Ribosomal proteins	0,36	2,77777778	-1,474	3,5353E-08	4,7166E-07
rplC	50S ribosomal protein L3	3.7.1 Ribosomal proteins	0,339	2,94985251	-1,561	1,1246E-09	2,0486E-08
rplD	50S ribosomal protein L4	3.7.1 Ribosomal proteins	0,358	2,79329609	-1,481	7,9489E-09	1,1913E-07
rplW	50S ribosomal protein L23	3.7.1 Ribosomal proteins	0,265	3,77358491	-1,914	1,0791E-14	4,5404E-13
rplB	50S ribosomal protein L2	3.7.1 Ribosomal proteins	0,193	5,18134715	-2,37	2,124E-15	1,0562E-13
rpsS	30S ribosomal protein S19	3.7.1 Ribosomal proteins	0,178	5,61797753	-2,493	2,4001E-18	2,0198E-16
rplV	50S ribosomal protein L22	3.7.1 Ribosomal proteins	0,2	5	-2,325	4,933E-11	1,1606E-09
rpsC	30S ribosomal protein S3	3.7.1 Ribosomal proteins	0,18	5,55555556	-2,47	8,7151E-12	2,2974E-10
rplP	50S ribosomal protein L16	3.7.1 Ribosomal proteins	0,135	7,40740741	-2,892	7,5514E-14	2,8487E-12
rpmC	50S ribosomal protein L29	3.7.1 Ribosomal proteins	0,131	7,63358779	-2,929	2,4445E-13	8,7682E-12
rpsQ	30S ribosomal protein S17	3.7.1 Ribosomal proteins	0,165	6,06060606	-2,601	1,9039E-10	4,1245E-09
rplN	50S ribosomal protein L14	3.7.1 Ribosomal proteins	0,203	4,92610837	-2,303	3,4743E-10	7,0388E-09
rplX	50S ribosomal protein L24	3.7.1 Ribosomal proteins	0,19	5,26315789	-2,393	4,2376E-10	8,4289E-09
rplE	50S ribosomal protein L5	3.7.1 Ribosomal proteins	0,183	5,46448087	-2,454	2,2501E-09	3,8165E-08
rpsN1, rpsZ	30S ribosomal protein S14 type Z	3.7.1 Ribosomal proteins	0,145	6,89655172	-2,783	1,3472E-08	1,9393E-07
rpsH	30S ribosomal protein S8	3.7.1 Ribosomal proteins	0,097	10,3092784	-3,358	1,557E-15	7,9444E-14
rplF	50S ribosomal protein L6	3.7.1 Ribosomal proteins	0,122	8,19672131	-3,032	1,5666E-12	5,0407E-11

rplR	ribosomal protein L18	3.7.1 Ribosomal proteins	0,113	8,84955752	-3,141	7,6417E-12	2,0642E-10
rpsE	30S ribosomal protein S5	3.7.1 Ribosomal proteins	0,164	6,09756098	-2,608	5,3134E-08	6,7201E-07
rpmD	50S ribosomal protein L30	3.7.1 Ribosomal proteins	0,132	7,57575758	-2,926	2,1034E-09	3,5955E-08
rplO	50S ribosomal protein L15	3.7.1 Ribosomal proteins	0,101	9,9009901	-3,305	9,1233E-12	2,3764E-10
rpsD	30S ribosomal protein S4	3.7.1 Ribosomal proteins	0,346	2,89017341	-1,531	3,8986E-06	3,2189E-05
rplQ	50S ribosomal protein L17	3.7.1 Ribosomal proteins	0,24	4,16666667	-2,057	8,8665E-11	2,0208E-09
rspl	30S ribosomal protein S9	3.7.1 Ribosomal proteins	0,629	1,58982512	-0,668	0,01958507	0,04740248
tyrS	Tyrosyl-tRNA synthetase (Tyrosine--tRNA ligase) (TyrRS)	3.7.2 Aminoacyl-tRNA synthetases	0,551	1,81488203	-0,86	0,00078642	0,00301876
pheT	Phenylalanyl-tRNA synthetase beta chain (Phenylalanine-- tRNA ligase beta chain) (PheRS)	3.7.2 Aminoacyl-tRNA synthetases	0,572	1,74825175	-0,805	0,00659853	0,01919891
proS	Prolyl-tRNA synthetase (Proline--tRNA ligase) (ProRS)	3.7.2 Aminoacyl-tRNA synthetases	0,622	1,60771704	-0,684	0,01208965	0,03202441
thrS	Threonyl-tRNA synthetase (Threonine--tRNA ligase) (ThrRS)	3.7.2 Aminoacyl-tRNA synthetases	0,303	3,30033003	-1,723	1,5531E-11	3,9978E-10
glnS	Glutamyl-tRNA synthetase	3.7.2 Aminoacyl-tRNA synthetases	0,421	2,37529691	-1,247	3,9158E-07	4,0799E-06
hisS	Histidyl-tRNA synthetase (Histidine--tRNA ligase) (HisRS)	3.7.2 Aminoacyl-tRNA synthetases	0,499	2,00400802	-1,004	3,7977E-05	0,00022158
aspS	Aspartyl-tRNA synthetase (Aspartate--tRNA ligase) (AspRS)	3.7.2 Aminoacyl-tRNA synthetases	0,469	2,13219616	-1,093	1,3341E-05	9,2323E-05
serS1	Seryl-tRNA synthetase	3.7.2 Aminoacyl-tRNA synthetases	0,398	2,51256281	-1,331	1,3601E-06	1,261E-05
ileS	Isoleucyl-tRNA synthetase (Isoleucine--tRNA ligase)	3.7.2 Aminoacyl-tRNA synthetases	0,327	3,05810398	-1,613	5,0045E-09	7,8775E-08
argS	Arginyl-tRNA synthetase (Arginine--tRNA ligase) (ArgRS)	3.7.2 Aminoacyl-tRNA synthetases	0,614	1,6286645	-0,705	0,01027066	0,02781214
gltX	Glutamyl-tRNA synthetase (Glutamate--tRNA ligase) (GluRS)	3.7.2 Aminoacyl-tRNA synthetases	0,543	1,84162063	-0,881	0,00044697	0,00187712

glyQS	Glycyl-tRNA synthetase (Glycine--tRNA ligase) (GlyRS)	3.7.2 Aminoacyl-tRNA synthetases	0,263	3,80228137	-1,926	1,9314E-09	3,3275E-08
trpS	Tryptophanyl-tRNA synthetase	3.7.2 Aminoacyl-tRNA synthetases	0,601	1,66389351	-0,733	0,01114231	0,02995009
metG	Methionyl-tRNA synthetase (Methionine--tRNA ligase) (MetRS)	3.7.2 Aminoacyl-tRNA synthetases	0,508	1,96850394	-0,977	0,00013296	0,0006642
valS	Valyl-tRNA synthetase (Valine--tRNA ligase) (ValRS)	3.7.2 Aminoacyl-tRNA synthetases	0,466	2,14592275	-1,101	4,3426E-05	0,00025136
infC	Translation initiation factor IF-3	3.7.3 Translation initiation	0,505	1,98019802	-0,985	0,00225997	0,00766639
fusA	Elongation factor G (EF-G)	3.7.4 Translation elongation	0,271	3,6900369	-1,886	4,9822E-10	9,6469E-09
tuf	Elongation factor Tu	3.7.4 Translation elongation	0,547	1,82815356	-0,87	0,00731936	0,02098919
frr	Ribosome recycling factor (Ribosome-releasing factor) (RRF)	3.7.5 Translation termination	0,514	1,94552529	-0,961	0,00035543	0,00155682
hflC1	Stomatin/prohibitin homolog	3.8 Protein modification	0,437	2,28832952	-1,193	6,5556E-07	6,4611E-06
PFCIRM1 38_00200	dolichyl-phosphate-mannose-protein mannosyltransferase (Glycosyl transferase, family 39)	3.8 Protein modification	0,455	2,1978022	-1,135	8,3235E-06	6,2799E-05
pknB	Serine/threonine protein kinase	3.8 Protein modification	0,369	2,7100271	-1,44	2,0423E-07	2,3202E-06
trxA2	Thioredoxin	3.8 Protein modification	0,165	6,06060606	-2,602	2,3871E-06	2,0726E-05
fkbP	FK506-binding protein (peptidyl-prolyl cis-trans isomerase)	3.8 Protein modification	0,457	2,18818381	-1,129	2,269E-05	0,00014348
ffh	GTP binding signal recognition particle protein	3.9 Protein folding	0,677	1,47710487	-0,564	0,01935583	0,04689984
groL2 (groEL2) (hsp60)	60 kDa chaperonin 2 (Protein Cpn60 2) (groEL protein 2) (Heat shock protein 60 2)	3.9 Protein folding	0,486	2,05761317	-1,041	0,00348033	0,01090968
groS1(gro ES1) (hsp10)	10 kDa chaperonin 1 (Protein Cpn10 1) (groES protein 1) (Heat shock 10 1)	3.9 Protein folding	0,394	2,53807107	-1,342	0,00049046	0,00202096
groL1 (groEL1)	60 kDa chaperonin 1 (Protein Cpn60 1) (groEL protein 1) (Heat shock protein 60 1)	3.9 Protein folding	0,304	3,28947368	-1,72	2,8762E-06	2,4504E-05

groS2 (groES2) (hsp10)	10 kDa chaperonin 2 (Protein Cpn10 2) (groES protein 2) (Heat shock protein 10 2)	3.9 Protein folding	0,603	1,65837479	-0,73	0,00968977	0,02650153
tig	Trigger factor (TF)	3.9 Protein folding	0,527	1,89753321	-0,924	0,00197876	0,0067649
PFCIRM1 38_01290	Beta-lactamase-like:RNA-metabolising metallo-beta-lactamase	4.2 Detoxification	0,39	2,56410256	-1,36	3,1801E-07	3,3614E-06
PFCIRM1 38_07575	beta-lactamase	4.2 Detoxification	0,53	1,88679245	-0,916	0,00488089	0,01476462
pcaC2	4-carboxymuconolactone decarboxylase	4.2 Detoxification	0,613	1,63132137	-0,706	0,0153359	0,03924556
casC	CRISPR system CASCADE complex protein CasC	4.4 Phage-related function	0,573	1,7452007	-0,803	0,0184017	0,04523923
casD	CRISPR system CASCADE complex protein CasD	4.4 Phage-related function	0,526	1,90114068	-0,926	0,00280899	0,00922833
casE	CRISPR system CASCADE complex protein CasE	4.4 Phage-related function	0,504	1,98412698	-0,989	0,00098685	0,00370365
cas1	CRISPR-associated endonuclease Cas1, ECOLI subtype	4.4 Phage-related function	0,469	2,13219616	-1,092	0,00134929	0,00485328
PFCIRM1 38_02315	Kinase	4.6 Miscellaneous	0,383	2,61096606	-1,385	0,00139952	0,00500353
PFCIRM1 38_02865	HIT family protein	4.6 Miscellaneous	0,415	2,40963855	-1,27	0,00143203	0,00511138
PFCIRM1 38_03220	GCN5-related N-acetyltransferase	4.6 Miscellaneous	0,54	1,85185185	-0,889	0,00174795	0,00611924
luxS	S-ribosylhomocysteine lyase (Autoinducer-2 production protein luxS) (AI-2 synthesis protein)	4.6 Miscellaneous	0,614	1,6286645	-0,705	0,02050961	0,04931322
oxaA	Membrane protein insertase	4.6 Miscellaneous	0,374	2,67379679	-1,418	2,1836E-07	2,4253E-06
npdA	Silent information regulator protein Sir2 /NAD-dependent deacetylase	4.6 Miscellaneous	0,446	2,24215247	-1,166	4,0683E-05	0,00023611
PFCIRM1 38_05780	Fe-S oxidoreductase	4.6 Miscellaneous	0,127	7,87401575	-2,974	3,6151E-16	2,0281E-14
PFCIRM1 38_05925	Oxidoreductase	4.6 Miscellaneous	0,153	6,53594771	-2,708	1,3825E-15	7,3778E-14
mviN	Conserved membrane protein, MviN-like protein	4.6 Miscellaneous	0,461	2,1691974	-1,118	0,0018332	0,00634657
lipA	Lipoic acid synthetase	4.6 Miscellaneous	0,564	1,77304965	-0,826	0,0006516	0,00258279

PFCIRM1 38_09940	Von Willebrand factor, type A	4.6 Miscellaneous	0,325	3,07692308	-1,621	4,4755E-06	3,6268E-05
PFCIRM1 38_10155	PP2Cc, Serine/threonine phosphatases,	4.6 Miscellaneous	0,627	1,59489633	-0,674	0,0196066	0,04740248
PFCIRM1 38_10615	dehydrogenase/reductase	4.6 Miscellaneous	0,661	1,5128593	-0,597	0,01863681	0,04558667
PFCIRM1 38_11500	Glycosyltransferase	4.6 Miscellaneous	0,52	1,92307692	-0,944	0,00117522	0,00431439
PFCIRM1 38_11855	oxidoreductase	4.6 Miscellaneous	0,43	2,3255814	-1,218	3,2795E-05	0,00019605
PFCIRM1 38_12335	protein of unknown function	5.1 Protein of unknown function similar to other proteins from the same organism	0,402	2,48756219	-1,315	9,4569E-06	6,9669E-05
PFCIRM1 38_00015	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,591	1,69204738	-0,759	0,00762589	0,0217451
PFCIRM1 38_00185	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,355	2,81690141	-1,494	3,2154E-09	5,173E-08
PFCIRM1 38_00660	Hypothetical membrane protein	5.2 Protein of unknown function similar to proteins from other organisms	0,492	2,03252033	-1,024	0,0004143	0,00176018
PFCIRM1 38_01150	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,559	1,78890877	-0,839	0,01546312	0,03952488
PFCIRM1	Hypothetical protein	5.2 Protein of	0,494	2,0242915	-1,018	0,00017924	0,00084337

38_01300		unknown function similar to proteins from other organisms					
PFCIRM1 38_01745	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,35	2,85714286	-1,517	2,9818E-06	2,5287E-05
PFCIRM1 38_01750	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,309	3,23624595	-1,692	3,4592E-05	0,00020456
PFCIRM1 38_01755	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,446	2,24215247	-1,164	1,472E-05	0,00010002
PFCIRM1 38_01760	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,523	1,91204589	-0,935	0,01895944	0,04609251
PFCIRM1 38_02105	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,618	1,61812298	-0,695	0,01661365	0,04187864
PFCIRM1 38_02305	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,35	2,85714286	-1,515	0,0028286	0,00927883
PFCIRM1 38_02320	UPF0210 protein	5.2 Protein of unknown function	0,415	2,40963855	-1,27	2,0854E-07	2,3399E-06



		similar to proteins from other organisms					
PFCIRM1 38_02810	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,484	2,0661157	-1,047	0,00058972	0,0023719
PFCIRM1 38_02830	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,493	2,02839757	-1,021	0,00122682	0,00446635
PFCIRM1 38_02950	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,52	1,92307692	-0,942	0,00133948	0,00482829
PFCIRM1 38_03420	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,5	2	-1,001	0,00015646	0,00075568
PFCIRM1 38_03690	Hypothetical transmembrane protein	5.2 Protein of unknown function similar to proteins from other organisms	0,553	1,80831826	-0,854	0,00030891	0,00137099
PFCIRM1 38_03855	Hypothetical membrane protein	5.2 Protein of unknown function similar to proteins from other organisms	0,621	1,61030596	-0,686	0,01606597	0,04077997
PFCIRM1 38_05525	Hypothetical membrane protein	5.2 Protein of unknown function similar to proteins	0,419	2,38663484	-1,256	1,5751E-06	1,436E-05

		from other organisms					
PFCIRM1 38_05935	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,179	5,58659218	-2,484	2,1105E-11	5,1884E-10
PFCIRM1 38_06015	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,458	2,18340611	-1,126	0,00014176	0,000697
PFCIRM1 38_06085	Putative uncharacterized protein	5.2 Protein of unknown function similar to proteins from other organisms	0,584	1,71232877	-0,776	0,00930584	0,02561156
PFCIRM1 38_07050	Hypothetical membrane protein	5.2 Protein of unknown function similar to proteins from other organisms	0,595	1,68067227	-0,748	0,00507038	0,01525997
PFCIRM1 38_07405	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,284	3,52112676	-1,818	0,00650729	0,01900928
PFCIRM1 38_07760	Hypothetical membrane protein	5.2 Protein of unknown function similar to proteins from other organisms	0,458	2,18340611	-1,127	0,00012091	0,00061098
PFCIRM1 38_08100	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other	0,55	1,81818182	-0,862	0,00227597	0,0076968

		organisms					
PFCIRM1 38_08555	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,255	3,92156863	-1,972	0,00065493	0,00259129
PFCIRM1 38_08630	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,566	1,76678445	-0,82	0,01379083	0,0358365
PFCIRM1 38_08980	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,362	2,76243094	-1,466	7,0988E-08	8,629E-07
PFCIRM1 38_10000	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,55	1,81818182	-0,863	0,00155047	0,00548049
PFCIRM1 38_10235	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,253	3,95256917	-1,98	1,8996E-07	2,1827E-06
PFCIRM1 38_10240	Hypothetical secreted and membrane protein	5.2 Protein of unknown function similar to proteins from other organisms	0,244	4,09836066	-2,037	7,8508E-12	2,0948E-10
PFCIRM1 38_10400	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,499	2,00400802	-1,004	0,00175393	0,00613034

PFCIRM1 38_10520	Hypothetical secreted protein	5.2 Protein of unknown function similar to proteins from other organisms	0,445	2,24719101	-1,167	0,00024499	0,00110979
PFCIRM1 38_10550	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,592	1,68918919	-0,757	0,00596512	0,01758988
PFCIRM1 38_10765	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,614	1,6286645	-0,703	0,01133687	0,03036117
PFCIRM1 38_11190	Hypothetical membrane protein	5.2 Protein of unknown function similar to proteins from other organisms	0,554	1,80505415	-0,851	0,00092164	0,00347082
PFCIRM1 38_11325	Hypothetical secreted protein	5.2 Protein of unknown function similar to proteins from other organisms	0,599	1,66944908	-0,74	0,00274834	0,00909738
PFCIRM1 38_11530	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,45	2,22222222	-1,152	0,00015461	0,00075007
PFCIRM1 38_11550	Hypothetical protein	5.2 Protein of unknown function similar to proteins from other organisms	0,602	1,66112957	-0,731	0,00647959	0,01895366
PFCIRM1	Hypothetical protein	5.2 Protein of	0,484	2,0661157	-1,046	0,00181393	0,0062998

38_11810		unknown function similar to proteins from other organisms					
PFCIRM1 38_12275	Hypothetical secreted protein	5.2 Protein of unknown function similar to proteins from other organisms	0,603	1,65837479	-0,73	0,00454104	0,01387682
PFCIRM1 38_00435	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	0,429	2,33100233	-1,22	0,00023681	0,0010776
PFCIRM1 38_01880	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	0,646	1,54798762	-0,631	0,01081575	0,02914391
PFCIRM1 38_02445	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	0,57	1,75438596	-0,811	0,00335633	0,01059689
PFCIRM1 38_02765	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	0,286	3,4965035	-1,807	0,00293814	0,00953805
PFCIRM1 38_04485	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	0,498	2,00803213	-1,006	0,01671466	0,04208477
PFCIRM1 38_04545	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	0,445	2,24719101	-1,168	0,00026335	0,00118804
PFCIRM1 38_05150	Hypothetical protein	6 Protein of unknown function, without similarity to	0,594	1,68350168	-0,753	0,00693117	0,01995447

		other proteins					
PFCIRM1 38_05205	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	0,606	1,65016502	-0,721	0,01284854	0,03378918
PFCIRM1 38_05510	Protein of unknown function	6 Protein of unknown function, without similarity to other proteins	0,21	4,76190476	-2,252	0,00157402	0,00555478
PFCIRM1 38_06570	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	0,338	2,95857988	-1,567	1,0521E-05	7,6224E-05
PFCIRM1 38_06675	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	0,635	1,57480315	-0,656	0,00665084	0,01928747
PFCIRM1 38_06720	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	0,576	1,73611111	-0,797	0,00346043	0,01089446
PFCIRM1 38_07920	protein of unknown function	6 Protein of unknown function, without similarity to other proteins	0,579	1,72711572	-0,789	0,01375217	0,03577853
PFCIRM1 38_08640	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	0,564	1,77304965	-0,826	0,00444145	0,01364873
PFCIRM1 38_09065	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	0,318	3,14465409	-1,654	1,1699E-08	1,7179E-07
PFCIRM1 38_10065	Hypothetical protein	6 Protein of unknown function, without similarity to	0,345	2,89855072	-1,533	3,8435E-06	3,1855E-05

		other proteins					
PFCIRM1 38_11045	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	0,502	1,99203187	-0,994	5,059E-05	0,00028751
PFCIRM1 38_11050	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	0,422	2,36966825	-1,244	0,00038207	0,00165211
PFCIRM1 38_11490	Hypothetical membrane protein	6 Protein of unknown function, without similarity to other proteins	0,559	1,78890877	-0,839	0,00296424	0,00958036
PFCIRM1 38_12015	Hypothetical protein	6 Protein of unknown function, without similarity to other proteins	0,524	1,90839695	-0,933	0,00112543	0,00416657

**Table S3:** RT-qPCR results for validation of RNAseq results.

Rnaseq		RT-qPCR validation												
Induction (I) Repression (R)	Fold	Target	Sample	Mean Cq	Mean Efficiency Corrected Cq	Normalized Expression	Relative Normalized Expression	Fold	Pvalue	Significativity P value <0.05	Forward primer	Reverse primer	Reference	
		5930	1d post-inoculation	17.1	16.71	23.43029	1	1	N/A					
<b>R</b>	<b>6</b>	<b>5930</b>	<b>3d post-inoculation</b>	<b>22.41</b>	<b>21.91</b>	<b>5.22024</b>	<b>0.2228</b>	<b>-4.48836</b>	<b>0.021089</b>	Yes	AAGATCC TGCCGA CCTATGA	TCCACACG GAGTTGTA CGG	this paper	
		5930	9d post-inoculation	23.38	22.85	4.28595	0.18292	-5.46676	0.021204	Yes				
		7015	1d post-inoculation	24.8	24.9	0.08008	1	1	N/A					
<b>I</b>	<b>49.4</b>	<b>7015</b>	<b>3d post-inoculation</b>	<b>23.27</b>	<b>23.37</b>	<b>1.89254</b>	<b>23.63262</b>	<b>23.63262</b>	<b>0.000142</b>	Yes	ACCCCTA CGGAAC GCAAG	GAGCAATC TGGTTGAC GTTGTT	this paper	
		7015	9d post-inoculation	23.75	23.86	2.13738	26.69002	26.69002	0.027467	Yes				
		argG	1d post-inoculation	22.04	21.57	0.80639	1	1	N/A					
<b>R</b>	<b>3.1</b>	<b>argG</b>	<b>3d post-inoculation</b>	<b>26.44</b>	<b>25.88</b>	<b>0.33294</b>	<b>0.41287</b>	<b>-2.42205</b>	<b>0.007666</b>	Yes	GACCATC AACGGC AAGAAGT	TCGTAGG CGATGAAC AACAG	Dalmasso 2012 PlosOne	
		argG	9d post-inoculation	27.67	27.09	0.22701	0.28151	-3.55	0.040364	Yes				



			on					23					
		fabF	1d post-inoculation	19.78	19.59	3.18311	1	1	N/A				
<b>R</b>	<b>6.3</b>	<b>fabF</b>	<b>3d post-inoculation</b>	<b>25.9</b>	<b>25.66</b>	<b>0.38881</b>	<b>0.12215</b>	<b>-8.18678</b>	<b>0.025484</b>	Yes	GGTGCC CTCACAC CTCTC	GATGACG CCGATTCC ACT	this paper
		fabF	9d post-inoculation	26.25	26.01	0.48128	0.1512	-6.61386	0.03531	Yes			
		gadB	1d post-inoculation	27.55	28.26	0.00785	1	1	N/A				
<b>I</b>	<b>5</b>	<b>gadB</b>	<b>3d post-inoculation</b>	<b>28.96</b>	<b>29.71</b>	<b>0.02349</b>	<b>2.9926</b>	<b>2.9926</b>	<b>0.010501</b>	Yes	ATCTCGG AGGACC ACAAGG	GTACATGC CGGTGTA GGTGA	this paper
		gadB	9d post-inoculation	24.48	25.11	0.89795	114.41965	114.41965	0.037762	Yes			
		gtfF	1d post-inoculation	30.1	27	0.01875	1	1	N/A				
<b>I</b>	<b>6.1</b>	<b>gtfF</b>	<b>3d post-inoculation</b>	<b>31.57</b>	<b>28.31</b>	<b>0.06173</b>	<b>3.29223</b>	<b>3.29223</b>	<b>0.011346</b>	Yes	CGACGAT TCCGACA AGCCCG AATAC	TGATCTCG TCGCTGTC GAGGATC AC	Deutsch et al. AEM 2010
		gtfF	9d post-inoculation	35.06	31.45	0.01109	0.59131	-1.69116	0.486605	Not			
		iolT1	1d post-inoculation	27.33	27.76	0.01108	1	1	N/A				
<b>I</b>	<b>11</b>	<b>iolT1</b>	<b>3d post-inoculation</b>	<b>29.05</b>	<b>29.5</b>	<b>0.02699</b>	<b>2.43566</b>	<b>2.43566</b>	<b>0.014506</b>	Yes	GCTCGAT CCCATT CCACC	TGTTGGGT GTCTCGG GGT	this paper

		iolT1	9d post-inoculation	25.18	25.58	0.64701	58.38609	58.38609	0.019236	Yes			
		narH	1d post-inoculation	30.4	28.61	0.00613	1	1	N/A				
<b>I</b>	<b>22</b>	<b>narH</b>	<b>3d post-inoculation</b>	<b>31.71</b>	<b>29.84</b>	<b>0.02142</b>	<b>3.49599</b>	<b>3.49599</b>	<b>0.006567</b>	Yes	GTGGGACCTCATCTCCACCT	CGTACCA GACCATC GGCATC	this paper
		narH	9d post-inoculation	28.15	26.49	0.344	56.13897	56.13897	0.075888	Not			
		sdhC1	1d post-inoculation	19.17	19.49	3.42533	1	1	N/A				
<b>R</b>	<b>8.6</b>	<b>sdhC1</b>	<b>3d post-inoculation</b>	<b>25.26</b>	<b>25.68</b>	<b>0.38255</b>	<b>0.11168</b>	<b>-8.95383</b>	<b>0.000378</b>	Yes	CACATGGTCGGAAATCTCAA	AGATCCACAGACCACCCTCA	this paper
		sdhC1	9d post-inoculation	32.38	32.92	0.004	0.00117	-856.93026	N/A	Not			
		slgT	1d post-inoculation	22.8	21.9	0.64339	1	1	N/A				
<b>R</b>	<b>4</b>	<b>slgT</b>	<b>3d post-inoculation</b>	<b>28.76</b>	<b>27.62</b>	<b>0.09946</b>	<b>0.15458</b>	<b>-6.46897</b>	<b>0.000094</b>	Yes	TCGGATACTGGACCACCAA	ACGATGAA GGGGATG AACAG	this paper
		slgT	9d post-inoculation	33.38	32.06	0.00723	0.01124	-88.99659	0.00001	Yes			
		slpE F2R2	1d post-inoculation	27.71	26.9	0.02012	1	1	N/A				
<b>I</b>	<b>12.9</b>	<b>slpE F2R2</b>	<b>3d post-inoculation</b>	<b>28.3</b>	<b>27.47</b>	<b>0.1102</b>	<b>5.47841</b>	<b>5.47841</b>	<b>0.017765</b>	Yes	ATCGTGTGCGTTG	CATCCGA GAACGTCT	this paper

			ion								ATTGCCT	GACCC	
		slpE F2R2	9d post- inoculation	32.7 6	31.8	0.00867	0.43103	- 2.32 005	0.4845 85	Not			
		mmda	1d post- inoculation	17.6 7	16.84	21.49449	1	1	N/A				
<b>R</b>	<b>2.4</b>	<b>mmda</b>	<b>3d post- inoculation</b>	<b>21.9 3</b>	<b>20.91</b>	<b>10.46725</b>	<b>0.48697</b>	<b>- 2.05 35</b>	<b>0.0050 65</b>	Yes	CGCGTC GAGAAG CAACATT C	AACGAATG GGGATCG AGCAG	this paper
		mmda	9d post- inoculation	22.4 1	21.36	12.07668	0.56185	- 1.77 983	0.4506 15	Not			
		dadA2	1d post- inoculation	28.2	27.77	0.01098	1	1	N/A				
<b>I</b>	<b>13</b>	<b>dadA2</b>	<b>3d post- inoculation</b>	<b>30.5</b>	<b>30.03</b>	<b>0.0187</b>	<b>1.70259</b>	<b>1.70 259</b>	<b>0.1238 29</b>	Not	CAAGTCG GAGTTCG TCAAGG	GCCCTGTT CGTGCATC TC	this paper
		dadA2	9d post- inoculation	28.5 7	28.14	0.10988	10.00451	10.0 0451	0.0864 16	Not			
normalizati on genes		<b>sdaA</b>								stable	CGCGGA ACTGCTC AATATCT	CACCTCGT CCTTCGAG TAGC	Dalmasso 2012 PlosOne
		<b>eno1</b>								stable	TACGAGT TCGAGG GCAAGTC	AAACCAGC GGGTAGT CATTG	Dalmasso 2012 PlosOne
		<b>cstA</b>								stable	CCTACAC ACCGGA GGAAGA G	GGACCGT CCATTTGT TATGC	Dalmasso 2012 PlosOne

**CAPÍTULO 3: Discussão geral,  
conclusão e perspectivas**

**CHAPITRE 3 : Discussion générale,  
conclusion et perspectives**

## Discussão geral

O objetivo geral desta tese foi de melhorar a nossa compreensão dos mecanismos envolvidos no fenômeno sobrevivência a longo prazo em *P. freudenreichii*. Uma das principais características da espécie é, de fato, uma boa sobrevivência a longo prazo, e este é um critério favorável à sua utilização em vários setores da indústria, tais como: a produção de queijos PPC, especialmente o Emmental, um dos produtos mais consumidos na França (Dalmasso *et al.* 2012a), e o desenvolvimento de produtos probióticos com efeitos benéficos para a saúde humana (Cousin *et al.* 2012). A síntese de biomoléculas, tais como ácido propiônico e vitamina B12 (Van Wyk, Witthuhn and Britz 2011; Wang and Yang 2013) é uma outra utilização industrial importante para esta espécie. Mas esta última é realizada sob condições de cultura altamente controlados para manter uma atividade metabólica ótima e a problemática de sobrevivência a longo prazo não está no escopo destas aplicações.

Neste trabalho, mostramos, através da análise do perfil de expressão de genes e da produção e consumação de metabólitos, que *P. freudenreichii* é capaz de reduzir o seu metabolismo, colocar-se em um estado de dormência e utilizar vias alternativas de energia. Estes resultados serão discutidos nos parágrafos seguintes.

### ➤ **Caracterização fenotípica de *P. freudenreichii*:**

Estudos anteriores relatam que *P. freudenreichii* é uma bactéria versátil que sobrevive a vários estresses ambientais (Jan, Rouault and Maubois 2000; Dalmasso *et al.* 2012b; Saraoui *et al.* 2013b). Além disso, *P. freudenreichii* é capaz de utilizar diferentes fontes de carbono como substrato e pode armazenar certos compostos para a sobrevivência a longo prazo (Thierry *et al.* 2011). Para uma melhor compreensão desta bactéria, nós realizamos a caracterização fenotípica de 8 linhagens de *P. freudenreichii*, durante um período de carência nutricional. Sabendo-se que, em condições ótimas, esta bactéria tem um tempo de geração de 5h (Falentin *et al.* 2010a), assumiu-se que um período de 11 dias (com uma 1 amostragem por dia) durante o qual a bactéria está sujeita ao estresse nutricional permitiu separar e caracterizar bem as diferentes fases de crescimento, até uma fase de sobrevivência a longo prazo. Testes de sobrevivência a longo prazo já foram realizados em outras bactérias indicando uma caracterização num tempo, significativamente, mais longo, como por exemplo, Finkel e Kolter que consideraram uma duração de um ano para a avaliação da sobrevivência em *Escherichia coli* (Finkel 2006b). Com certeza seria interessante fazer análises mais longas, mas um

dos objetivos desta tese foi de tentar elucidar os mecanismos que permitiram a adaptação de *P. freudenreichii* à esta fase de crescimento.

A capacidade de sobrevivência a longo prazo varia em função do gênero, da espécie ou mesmo da linhagem estudada. Os estudos fenotípicos de 8 linhagens demonstraram que o conjunto de todas as linhagens de *P. freudenreichii* apresentaram uma boa sobrevivência a longo prazo, no entanto, com algumas variações entre as linhagens. Estudar este fenômeno durante um período de 6 ou 12 meses, nos teria permitido de determinar durante quanto tempo era possível a esta bactéria de sobreviver em carência nutricional e, sobretudo, poderia ter revelado e ou enfatizado as diferenças entre as linhagens durante a sobrevivência a longo prazo. Esta capacidade poderia de fato tornar-se um critério para a seleção de linhagens para aplicações em queijos ou probióticas de *P. freudenreichii*.

➤ **Dormência em *P. freudenreichii*:**

Muitas bactérias são capazes de entrar em um estado de dormência permitindo uma maior resistência e uma melhor sobrevivência a ambientes desfavoráveis (Rittershaus, Baek and Sasseti 2013b). Esta adaptação é caracterizada pela manutenção da viabilidade, no entanto, associada a uma perda de cultivabilidade em condições padrão laboratoriais. Este estado também é descrito pelo termo de estado VBNC (*Viable But Non Culturable*). O estado VBNC constitui uma estratégia muito importante para algumas espécies bacterianas, Gram positivas ou negativas, para a sobrevivência durante períodos prolongados de carência. Muitos estudos científicos têm sido dedicados a este tema. Uma busca PubMed mostra que o termo "viable but nonculturable bacteria" totaliza em 385 resultados. Dentre eles, 64 artigos sobre bactérias patogênicas ("VBNC" AND "pathogen") e 78 artigos sobre bactérias em estado VBNC em alimentos ou na água ("VBNC" AND "food"). Nesta última pesquisa, a maioria dos artigos tratavam-se patógenos alimentares (*Campylobacter* sp., *E. coli*, *Vibrio* sp., *Salmonella* sp.) e alguns poucos, eventualmente, sobre as bactérias de interesse alimentar (*food grade*). Três artigos relacionados com probióticos. Esta diferença de tratamento ou de interesse entre estas áreas é em parte explicado pelo fato de que muitas bactérias patogênicas são capazes de induzir a mesma infecção ou intoxicação mesmo estando em um estado VBNC, e geralmente, essas bactérias (não cultiváveis) não são detectadas por técnicas microbiológicas convencionais, como a contagem das UFC. Os estudos sobre probióticos são realizados para desenvolver a sua

importância na formulação destes produtos, uma vez que o número de bactérias viáveis no produto final é um critério levado em consideração para os efeitos probióticos sobre o hospedeiro. Em um estudo realizado por Davis e colaboradores, demonstrou-se que as bactérias que haviam propriedades probióticas entravam em estado VBNC e que elas só podiam ser quantificadas corretamente através de métodos alternativos, tais como hibridação *in situ* fluorescente, qPCR, RT-PCR, propídio monoazida-PCR e/ou a separação de células ativadas por fluorescência (Davis 2014). Desta mesma maneira, confirmamos o estado VBNC em *P. freudenreichii* CIRM-BIA138, linhagem mais resistente a carência nutricional, combinando diferentes metodologias, tais como a contagem de UFC, a microscopia de epi-fluorescência e qPCR.

➤ **Genes diferencialmente expressos em fases de crescimento exponencial e estacionária de *P. freudenreichii*:**

Para aprofundar os estudos relacionados com a sobrevivência a longo prazo desta bactéria, foi realizada a análise do perfil transcricional da linhagem *P. freudenreichii* CIRM- BIA138 utilizando a técnica de RNA-Seq, durante um período de carência nutricional. Inicialmente, escolhemos três pontos de crescimento bacteriano para realizar esta análise: exponencial, estacionária e estacionária tardia. Para obter estas amostras, realizamos a cultura bacteriana, a extração de RNA total e a depleção de RNAr (RNA ribossômico). Em seguida, sequenciamos o cDNA destas amostras pela plataforma Illumina HiSeq 2000, uma tecnologia de sequenciamento de próxima geração. Infelizmente, observa-se que não foi possível obter bons resultados de sequenciamento nas amostras da fase estacionária tardia. Isto, porque os resultados das 3 replicatas desta última fase apresentaram um perfil de contaminação por DNA genômico, uma vez que as *reads* estavam alinhadas de maneira uniforme e não-orientada, impedindo assim a interpretação dos dados de sequenciamento. No entanto, a análise de RNA-Seq entre as fases de crescimento bacteriano exponencial e estacionária nos permitiu identificar 912 genes diferencialmente expressos e potencialmente envolvidos na preparação e adaptação da linhagem à carência nutricional.

Atualmente, um dos poucos estudos de transcriptoma sobre o tema de sobrevivência a longo prazo é o realizado sobre o patógeno *Listeria monocytogenes* que provoca infecções de origem alimentar. A cultura de *L. monocytogenes* foi amostrada em diferentes fases do desenvolvimento (13h, 17h, 24h e 168h, 336h), num meio de cultura de triptona de soja ágar com extrato de levedura (Becton Dickinson, MD). Este

estudo foi realizado por meio de técnicas de microarranjo e RT-qPCR que revelaram genes induzidos envolvidos em: (i) síntese de aminoácidos, de ácidos graxos, de proteínas, de fosfolípidos, de envelope celular e de ribonucleótidos ; (ii) na transcrição; (iii) na desintoxicação; (iv) no transporte de proteínas; e (v) na divisão celular. Estes metabolismos altamente expressos durante a fase logarítmica são reprimidos durante o período de sobrevivência a longo prazo (Wen *et al.* 2011).

Observou-se ainda que os genes envolvidos na síntese de proteínas foram reprimidos, o que indica também, que esta bactéria, entra em estado de dormência após períodos prolongados (Wen *et al.* 2011). Apesar de *L. monocytogenes* ser patogênica, a estratégia de sobrevivência adotada é semelhante a adotada por *P. freudenreichii*. Os resultados do transcriptoma de *P. freudenreichii* CIRM-BIA138 também revelaram que, na fase estacionária, a bactérias começa a se preparar para um estado de dormência que lhe permite sobreviver e de utilizar fontes alternativas para obter e armazenar a energia.

Além disso, os nossos resultados demonstram uma redução no metabolismo de *P. freudenreichii*, uma vez que os genes envolvidos na fosforilação oxidativa e no ciclo de Wood-Werkman estão reprimidos durante a fase estacionária. Estes resultados estão de acordo com o trabalho de Saraoui e colaboradores que mostram que, em *P. freudenreichii* CIRM-BIA1, as vias da glicólise, do ciclo de Wood-Werkman e fosforilação oxidativa foram reprimidos em fase estacionária no cólon de suínos (Saraoui *et al.* 2013a).

➤ **Mecanismos moleculares envolvidos na sobrevivência a longo prazo de *P. freudenreichii*:**

Através de dados transcriptômicos e bioquímicos de *P. freudenreichii* CIRM-BIA138, obtivemos pistas que indicam quais genes estão ativos nesta bactéria e quais são os fatores ambientais envolvidos no desencadeamento de mecanismos moleculares que preparam a sobrevivência bacteriana a longo prazo, durante um período de carência nutricional.

Durante este período, a linhagem *P. freudenreichii* CIRM-BIA138 é também capaz de explorar novas fontes de energia, tais como aminoácidos (como o glutamato, glicina e alanina), a arabinose e o inositol que são fontes de carbono alternativas. Para equilibrar o seu metabolismo, na ausência de oxigênio, esta bactéria utiliza provavelmente o polifosfato e o pirofosfato para armazenar a energia, os quais os genes estão presentes no genoma desta espécie (Falentin *et al.* 2010a). Assim, neste trabalho,



os genes envolvidos na síntese de polifosfatos são induzidos durante a fase estacionária para a utilização de ATP. Além disso, nesta última fase, uma diminuição na atividade de transcrição, tradução e secreção de proteínas foi observada. Outros parâmetros experimentais também foram analisados neste estudo, como a dosagem de compostos bioquímicos, demonstrando a influência destes últimos sobre o metabolismo de *P. freudenreichii* CIRM-BIA138, durante a carência nutricional. Alterações na síntese e no consumo de ácidos orgânicos, de açúcares e de aminoácidos também foram observados. Isto demonstra a capacidade de *P. freudenreichii* CIRM-BIA138 se adaptar ao ambiente, e contribui para uma melhor compreensão do conhecimento biológico desta espécie.

Todos estes dados fenotípicos, transcriptômicos e metabolômicos de *P. freudenreichii* CIRM-BIA138 são de grande importância para a compreensão desta bactéria, especialmente, durante as fases importantes encontradas, como por exemplo, durante a maturação de queijos PPC. No entanto, nas condições de maturação, *P. freudenreichii* atinge rapidamente uma fase estacionária de crescimento, mas raramente associada a uma carência de carbono. Este tipo de trabalho já havia sido começado na equipe, portanto com análise transcricional por microarranjo para identificar estratégias de adaptação de *P. freudentreichii* nas condições de fabricação de queijos (Dalmaso *et al.* 2012a). Nossos resultados de RNA-Seq vieram para complementar estes dados. Além disso, nosso estudo metabolômico foi muito direcionado a determinados compostos. Seria interessante fazer um estudo metabolômico mais abrangentes como o trabalho recente de Le Boucher (Le Boucher *et al.* 2013). Este último trabalho demonstrou a capacidade da espectrometria de massa – *metabolic fingerprinting* – a detectar alterações induzidas pelo metabolismo bacteriano no queijo durante 0, 8 e 48 h. Este estudo encontrou variações em 45 metabólitos, tais como aminoácidos, vitaminas e metabólitos voláteis (Le Boucher *et al.* 2013).

A interdisciplinaridade e os diferentes métodos aplicados neste trabalho de tese permitiram de identificar mecanismos que poderiam explicar o fenômeno de sobrevivência a longo prazo nesta bactéria. Uma destas estratégias de sobrevivência parece ser o uso de vias alternativas de energia para manter a atividade metabólica. Nestas etapas, a bactéria se prepara a entrar em um estado de dormência para enfrentar um período de estresse nutricional e sobreviver a longo prazo, enquanto aguarda a ocorrência de condições mais favoráveis.

## Discussion générale

L'objectif général de cette thèse était d'améliorer notre compréhension des mécanismes impliqués dans le phénomène de survie à long terme chez *P. freudenreichii*. Une des principales caractéristiques de l'espèce est, en effet, sa bonne survie à long terme et ceci est un critère favorable à son utilisation dans divers secteurs de l'industrie tels que : la production des fromages PPC, en particulier l'Emmental, l'un des produits les plus consommés en France (Dalmasso *et al.* 2012a) et le développement de produits probiotiques ayant des effets bénéfiques sur la santé humaine (Cousin *et al.* 2012). La synthèse de biomolécules, comme l'acide propionique et la vitamine B12 (Van Wyk, Witthuhn and Britz 2011; Wang and Yang 2013) est une autre utilisation industrielle d'importance pour cette espèce. Mais celle-ci s'effectue dans des conditions de culture hautement contrôlées permettant de maintenir une activité métabolique optimale et la problématique de survie à long terme est de fait hors champs pour ces applications.

Dans ce travail, nous avons montré, à travers l'analyse du profil d'expression des gènes et de la production et consommation de métabolites, que *P. freudenreichii* est capable de réduire son métabolisme, de se mettre dans un état de dormance et d'utiliser des voies d'énergie alternatives. Ces résultats seront discutés dans les paragraphes suivants.

### ➤ **Caractérisation phénotypique de *P. freudenreichii*:**

Des études antérieures rapportent que *P. freudenreichii* est une bactérie versatile qui survit à divers stress environnementaux (Jan, Rouault and Maubois 2000; Dalmasso *et al.* 2012a; Saraoui *et al.* 2013a). En outre, *P. freudenreichii* est capable d'utiliser différentes sources de carbone comme substrat et peut stocker certains composés pour sa survie à long terme (Thierry *et al.* 2011). Aussi, pour une meilleure compréhension de cette bactérie, nous avons effectué la caractérisation phénotypique de 8 souches de *P. freudenreichii*, au cours d'une période de carence nutritionnelle. Sachant que, dans des conditions optimales, cette bactérie a un temps de génération de 5h (Falentin *et al.* 2010b), nous avons supposé qu'une période de 11 jours (avec un 1 échantillonnage par jour) durant laquelle la bactérie est soumise au stress nutritionnel permettait de séparer et bien caractériser les différentes phases de croissance, jusqu'à une phase de survie à long terme. Des essais de survie à long terme avaient déjà été réalisés chez d'autres bactéries indiquant une caractérisation sur un temps nettement plus long, comme par exemple, Finkel et Kolter qui ont considéré une durée d'un an pour l'évaluation de la

survie chez *Escherichia coli* (Finkel 2006b). Il serait sûrement intéressant de disposer d'essais plus longs, mais un des buts de cette thèse était de tenter d'élucider les mécanismes qui permettait à l'adaptation de *P. freudenreichii* à cette phase de croissance.

La capacité de survie à long terme varie en fonction du genre, de l'espèce, voire de la souche étudiée. Les études phénotypiques sur 8 souches ont montré que l'ensemble des souches de *P. freudenreichii* présentait une bonne survie à long terme avec cependant quelques variations inter-souches. Etudier ce phénomène sur une durée de 6 ou 12 mois, nous aurait permis de déterminer pendant combien de temps encore il était possible à cette bactérie de survivre en carence nutritionnelle et, surtout, aurait peut-être permis de révéler et ou accentuer des différences inter-souches en matière de survie à long terme. Cette capacité pourrait en effet devenir un critère de sélection de souches pour les applications fromagères ou probiotiques de *P. freudenreichii*.

#### ➤ **La dormance chez *P. freudenreichii* :**

De nombreuses bactéries sont capables d'entrer dans un état de dormance permettant une plus grande résistance et une meilleure survie à des environnements défavorables (Rittershaus, Baek and Sasseti 2013b). Cette adaptation se caractérise par le maintien de la viabilité associée cependant à une perte de cultivabilité en conditions de laboratoire classiques. Cet état est également décrit sous le terme d'état VBNC (pour Viable But Non Culturable). L'état VBNC constitue une stratégie très importante chez certaines espèces bactériennes, à Gram positif ou négatif, pour la survie pendant de longues périodes de carence. De nombreuses études scientifiques lui ont été consacrées. Une recherche PubMed montre que le terme « viable but nonculturable bacteria » totalise 385 résultats. Parmi eux, 64 articles concernent les bactéries pathogènes (« VBNC » AND « pathogen ») et 78 articles portent sur les bactéries en état VBNC dans l'aliment ou l'eau (« VBNC » AND « food »). Dans cette dernière recherche, la majorité des articles ont trait aux pathogènes alimentaires (*Campylobacter* sp., *E. coli*, *Vibrio* sp., *Salmonella* sp.) et très peu, finalement, portent sur des bactéries d'intérêt alimentaire (food grade). Trois articles ont trait aux probiotiques. Cette différence de traitement ou d'intérêt entre ces divers domaines s'explique entre autres par le fait que beaucoup de bactéries pathogènes sont capables d'induire une infection ou une intoxication même à partir d'un état VBNC, et généralement ces bactéries (non cultivables) ne sont pas détectées par des techniques de microbiologie classique, telles

que le comptage des UFC. Les études portant sur les probiotiques sont amenées à se développer de par leur importance pour la formulation de ces produits, car le nombre de bactéries viables dans le produit final est un critère pris en compte pour l'effet probiotique sur l'hôte. Dans une étude de Davis et collaborateurs, il a été démontré que les bactéries ayant des propriétés probiotiques se mettaient en état VBNC et qu'elles ne pouvaient être quantifiées correctement qu'à travers des méthodes alternatives comme l'hybridation *in situ* en fluorescence, la PCRq, la RT-PCR, le propidium monoazide-PCR et/ou la séparation des cellules activées par fluorescence (Davis 2014). De même, nous avons confirmé l'état VBNC chez *P. freudenreichii* CIRM-BIA138, la souche la plus résistante à la carence nutritionnelle, en combinant différentes méthodologies telles que la numération des UFC, la microscopie à épi-fluorescence et la PCRq.

➤ **Gènes différentiellement exprimés dans les phases de croissance exponentielle et stationnaire de *P. freudenreichii* :**

Pour approfondir les études liées à la survie à long terme de cette bactérie, nous avons effectué l'analyse du profil transcriptionnel de la souche *P. freudenreichii* CIRM-BIA138 en utilisant la technique de RNA-Seq, au cours d'une période de carence nutritionnelle. Tout d'abord, nous avons choisi trois points de la croissance bactérienne pour effectuer cette analyse : exponentiel, stationnaire et stationnaire tardif. Pour obtenir ces échantillons, nous avons réalisé une culture bactérienne, une extraction des ARN totaux et une déplétion des ARNr (ARN ribosomiques). Ensuite, nous avons séquencé le cDNA de ces échantillons par la plateforme Illumina HiSeq 2000, une technologie de séquençage de nouvelle génération. Malheureusement, nous avons observé qu'il n'était pas possible d'obtenir de bons résultats de séquençage sur les échantillons de la phase stationnaire tardive. En effet, les résultats des 3 replicates de cette dernière phase ont présenté un profil de contamination par de l'ADN génomique, car les *reads* se sont alignés de manière uniforme et non orientée, empêchant ainsi l'interprétation des résultats de séquençage. Cependant, l'analyse du RNA-Seq entre les phases de croissance bactérienne exponentielle et stationnaire nous a permis d'identifier 912 gènes différentiellement exprimés et potentiellement associés à la préparation et l'adaptation de la souche à la carence nutritionnelle.

Actuellement, l'une des rares études de transcriptome sur le thème de la survie à long terme est celle menée sur le pathogène *Listeria monocytogenes* qui provoque des

infections d'origine alimentaire. La culture de *L. monocytogenes* a été échantillonnée à différentes phases de développement (13h, 17h, 24h et 168h, 336h), dans un milieu de culture gélose tryptone soja avec de l'extrait de levure (Becton Dickinson, MD). Cette étude a été réalisée à travers des techniques utilisant des puces à ADN et la RT-qPCR qui ont révélé des gènes inductibles impliqués dans : (i) la synthèse d'acides aminés, d'acides gras, de protéines, de phospholipides, de l'enveloppe cellulaire et de ribonucléotides; (ii) la transcription ; (iii) la détoxification ; (iv) le transport de protéines ; et (v) la division cellulaire. Ces métabolismes fortement exprimés pendant la phase logarithmique sont réprimés pendant la période de survie à long terme (Wen *et al.* 2011).

Il a également été observé que les gènes impliqués dans la synthèse des protéines ont été réprimés, ce qui indique aussi, chez cette bactérie, l'entrée en état de dormance après des périodes prolongées (Wen *et al.* 2011). Bien que *L. monocytogenes* soit pathogène, la stratégie de survie qu'elle adopte est similaire à celle adoptée par *P. freudenreichii*. Les résultats du transcriptome de *P. freudenreichii* CIRM-BIA138 ont également révélé que, dans la phase stationnaire, la bactérie commence à se préparer pour un état de dormance qui lui permet de survivre et d'utiliser des sources alternatives pour obtenir et stocker de l'énergie.

En outre, nos résultats démontrent une réduction du métabolisme de *P. freudenreichii*, puisque les gènes impliqués dans la phosphorylation oxydative et le cycle Wood-Werkman sont réprimés pendant la phase stationnaire. Ces résultats sont en accord avec le travail de Saraoui et collaborateurs qui montrent que, chez *P. freudenreichii* CIRM-BIA1, les voies de glycolyse, du cycle de Wood-Werkman et de la phosphorylation oxydative étaient régulé négativement en phase stationnaire dans un côlon de cochons (Saraoui *et al.* 2013a).

➤ **Mécanismes moléculaires impliqués dans la survie à long terme de *P. freudenreichii* :**

À travers des données transcriptomiques et biochimiques de *P. freudenreichii* CIRM-BIA138, nous avons obtenu des pistes indiquant quels gènes sont actifs dans cette bactérie et quels sont les facteurs environnementaux impliqués dans le déclenchement des mécanismes moléculaires qui préparent la survie bactérienne à long terme, au cours d'une période de carence nutritionnelle.

Au cours de cette période, la souche *P. freudenreichii* CIRM-BIA138 est aussi capable d'explorer de nouvelles sources d'énergie, telles que des acides aminés (comme le glutamate, la glycine et l'alanine), l'arabinose et l'inositol qui sont des sources de carbone alternatives. Pour équilibrer son métabolisme, en l'absence d'oxygène, cette bactérie utilise probablement le polyphosphate et pyrophosphate pour stocker l'énergie et qui les gènes sont présentes dans le génome de cette espèce (Falentin *et al.* 2010b). Ainsi dans ce travail, les gènes impliqués dans la synthèse de polyphosphates sont induits au cours de la phase stationnaire pour utiliser l'ATP. En outre, dans cette dernière phase, une diminution de l'activité de la transcription, de la traduction et de la sécrétion des protéines a été observée. D'autres paramètres expérimentaux ont aussi été analysés dans ce travail, comme le dosage des composés biochimiques, montrant l'influence de ces derniers sur le métabolisme de *P. freudenreichii* CIRM-BIA138, au cours de la carence nutritionnelle. Des altérations dans la synthèse et la consommation d'acides organiques, de sucres et d'acides aminés ont également été observés. Cela démontre la capacité de *P. freudenreichii* CIRM-BIA138 à s'adapter à l'environnement, et contribue à une meilleure compréhension de la connaissance biologique de cette espèce.

Toutes ces données phénotypiques, transcriptomiques et métabolomiques de *P. freudenreichii* CIRM-BIA138 sont d'une grande importance pour la compréhension de cette bactérie notamment pendant des étapes importantes rencontrées, par exemple, au cours de l'affinage des fromages PPC. Toutefois, dans les conditions d'affinage, *P. freudenreichii* atteint rapidement une phase stationnaire de croissance mais rarement liée à une carence carbone. Ce type de travail avait déjà commencé dans l'équipe lors de l'analyse transcriptionnelle par puce à ADN pour identifier les stratégies d'adaptation de *P. freudentreichii* dans les conditions de fabrication fromagère (Dalmaso *et al.* 2012a). Nos résultats RNA-Seq viennent compléter ces données. En outre, notre étude métabolomique était très ciblée sur certains composés. Il serait intéressant de faire une étude métabolomique plus exhaustive à l'instar des travaux récemment menés par Le Boucher (Le Boucher *et al.* 2013). Ce dernier travail a montré la capacité de la spectrométrie de masse - *metabolic fingerprinting* – à détecter les modifications induites par le métabolisme bactérien dans le fromage au cours de 0, 8 et 48 h. Cette étude a révélé des variations sur 45 métabolites, tels que les acides aminés, les métabolites volatils et les vitamines (Le Boucher *et al.* 2013).

L'interdisciplinarité et les différentes méthodes appliquées dans ce travail de thèse ont permis d'identifier des mécanismes qui pourraient expliquer le phénomène de survie à long terme de cette bactérie. Une de ces stratégies de survie semble être l'utilisation de voies énergétiques alternatives permettant le maintien de l'activité métabolique. Dans ces étapes, la bactérie se prépare à entrer en dormance pour affronter une période de stress nutritionnel et survivre à long terme, en attendant la survenue de conditions plus favorables.

## **Conclusão/ Conclusion**



## **Conclusão**

Neste presente trabalho de tese, apresentamos, pela primeira vez, a identificação de estratégias de sobrevivência a longo-prazo em *P. freudenreichii*, combinando as abordagens de microbiologia, genômica e transcriptômica. Desde a caracterização fenotípica à análise funcional do genoma de *P. freudenreichii*, nossos estudos demonstraram que esta bactéria diversifica suas fontes de energia, em seguida, diminuem o seu metabolismo energético para entrar em um estado de dormência, o qual elas são capazes de sobreviver por períodos prolongados. Além disso, este trabalho de tese fornece informações sobre a biologia do grupo Propionibacteria lacteos.

## **Conclusion**

Dans ce travail de thèse, nous avons présenté, pour la première fois, l'identification de stratégies de survie à long terme chez *P. freudenreichii*, en combinant des approches microbiologique, génomique et transcriptomique. De la caractérisation phénotypique à l'analyse fonctionnelle du génome de *P. freudenreichii*, nos études ont permis de démontrer que cette bactérie diversifie ses sources d'énergie, puis diminue son métabolisme énergétique pour entrer dans un état de dormance pendant lequel elle est capable de survivre durant de longues périodes. En outre, ce travail de thèse fournit des informations sur la biologie du groupe de Propionibacteries lactiques.

## **Perspectivas/ Perspectives**

## Perspectivas

Os resultados obtidos ao longo deste trabalho de tese nos permitiram a melhor compreender o fenômeno de sobrevivência a longo prazo em *P. freudenreichii*. Eles abrem várias perspectivas. De um modo, os dados gerados podem ser completados e explorados para melhorar o nosso conhecimento da genômica de *P. freudenreichii* (tópicos 1 e 2 abaixo). E de outro modo, ainda há muito a ser feito na análise de sobrevivência a longo prazo e diversas pistas poderiam ser utilizadas para completar este trabalho (tópicos 3-7).

### **(I) Sequência completa do genoma de *P. freudenreichii* CIRM-BIA138:**

Como uma primeira abordagem, podemos considerar que um genoma completo (isto é, sem *gaps*) não é essencial, se podemos extrair as informações necessárias das análises de transcriptômica e metabolômica desta bactéria. No entanto, um genoma completamente montado permite uma melhor anotação e, assim, nos dará uma melhor qualidade de informação genética sobre o organismo.

Para obter o genoma completo de *P. freudenreichii* CIRM-BIA138, uma opção seria realizar um mapeamento óptico do genoma desta linhagem. Isto permitiria a superação de dificuldades encontradas na etapa de montagem deste genoma, como por exemplo, a presença de uma grande quantidade de transposases. O mapeamento óptico de fato dá uma visão geral do genoma, auxiliando a reorganização de *contigs* e ao fechamento de *gaps*. Além disso, é possível combinar diferentes programas de montagem para melhorar o processo como um todo. De outra maneira, como já foi demonstrado em estudos precedentes (Garber *et al.* 2011), e para complementar as abordagens anteriores, é possível utilizar os dados do transcriptoma provenientes do RNA-Seq para alinhar os transcritos no genoma e fechar os *gaps*. Os dados de RNA-Seq também podem auxiliar no processo de anotação do genoma, utilizando programas como o ReadXplorer, que permite o alinhamento dos dados de expressão de cada gene sob o genoma para refinar a anotação (Hilker *et al.* 2014).

### **(II) Identificação de novos elementos no genoma de *P. freudenreichii* CIRM-BIA 138:**

Os dados de RNA-Seq permitem a descoberta de novos transcritos que não foram identificados e anotados em versões anteriores do genoma. Uma análise detalhada

permitirá também de identificar pequenos RNAs não-codificantes em *P. freudenreichii* CIRM-BIA138. A tecnologia de RNA-Seq revelou-se eficaz para a detecção de pequenos RNAs em *Pseudomonas aeruginosa* (Gómez-Lozano *et al.* 2012), e para a descoberta de novos transcritos em várias espécies bacterianas (McClure *et al.* 2013). Programas de bioinformática, como Rockhopper, através de uma interface gráfica amigável, permite que o usuário entre com genomas de referência e modifique as configurações, tais como: leituras *paired-end* ou *mate-pair*; alinhamento de leituras de várias replicatas; número de replicatas e condições experimentais diferentes utilizadas para a análise de RNA-Seq (McClure *et al.* 2013).

### **(III) Completação dos dados de RNA-Seq na fase estacionária tardia de *P. freudenreichii* CIRM-BIA138:**

Durante a análise de RNA-Seq da linhagem *P. freudenreichii* CIRM-BIA138, a quantidade de RNA extraída de amostras associadas com a fase de sobrevivência a longo prazo foi muito baixa para permitir um sequenciamento de qualidade. Além disso, as amostras da fase estacionária tardia foram contaminadas com DNA genômico. Isto limitou severamente os resultados obtidos. Por conseguinte, seria vantajoso obter uma maior quantidade de RNA e a melhorar a sua qualidade, aumentando o volume da cultura bacteriana amostrada (por exemplo 25 mL) visando melhorar, a partir disso, as próximas etapas de extração e de depleção dos rRNA certificando-se de remover o DNA genômico por tratamentos com DNase.

### **(IV) Mecanismo de dormência na espécie *P. freudenreichii*:**

Muitas bactérias Gram-positivas e Gram-negativas entram em dormência na fase de sobrevivência a longo prazo. Como vimos, dentre as 8 linhagens de *P. freudenreichii* testadas, todas sobrevivem bem e suas populações permanecem em níveis elevados durante o período de carência nutricional. No entanto, observamos que a linhagem CIRM-BIA138 apresenta um nível de sobrevivência maior do que outras linhagens. Esta linhagem foi objeto de uma análise mais aprofundada (demonstração do estado VBNC e RNA-Seq). Neste contexto, seria interessante complementar os resultados obtidos nesta tese, explorando a capacidade de sobrevivência a longo prazo e o estado VBNC em outras linhagens. Isso talvez permita identificar linhagens de *P. freudenreichii* com baixa sobrevivência a longo prazo, e em seguida, por genômica comparativa os genes envolvidos (por exemplo, pela sua ausência ou por serem

pseudogenes) nesta sobrevivência inferior. Do mesmo modo, outras formas de stress (por exemplo, calor, ácido ou osmótico) podem ser aplicados a culturas em estado de sobrevivência a longo prazo e carência nutricional. Isto permitiria a identificar os fenômenos de adaptação e resistência cruzada que poderiam nos esclarecer sobre os mecanismos implementados e nos fornecer uma melhor explicação do comportamento da espécie em condições de fabricação de queijos.

Além disso, este estudo pode ser complementado pela construção de mutantes visando a interromper os genes envolvidos no desencadeamento da dormência em *P. freudenreichii*. Trabalho abordando esta metodologia (construção de mutantes) já foi realizado para mostrar o envolvimento do gene *clpX* em *S. typhimurium* (Kusumoto, Miyashita and Kawamoto 2013) e *rpoS* em *E. coli* (Boaretti *et al.* 2003) na sobrevivência a longo prazo nestas espécies.

#### **(V) Reanimação na espécie *P. freudenreichii*:**

Ao sair do estado VBNC, muitas espécies bacterianas passam por um estado chamado de ressuscitação permitindo-lhes retomar o crescimento normal. A reanimação (ou ressuscitação), por vezes, requer condições especiais de crescimento. Da mesma forma, alguns genes parecem estar diretamente envolvidos neste fenômeno. Assim, em *M. luteus*, tem sido mostrado que as células são reanimadas depois da produção do fator Rpf (*Ressuscitation Promoting Factor*) (Greenblatt *et al.* 2004a). A investigação deste estado de reanimação nas células de *P. freudenreichii* seria de grande interesse para a conclusão deste trabalho. No entanto, é necessário realizar experimentos adicionais que comprovem que as células são ressuscitadas após o estado VBNC. Isto requer a identificação de condições de cultura que provocam a liberação do estado VBNC e da retomada do crescimento normal.

#### **(VI) Sobrevivência a longo prazo em *P. freudenreichii* CIRM-BIA138 na matriz de queijo:**

Os nossos resultados permitem uma melhor compreensão dos mecanismos moleculares envolvidos na sobrevivência a longo prazo de *P. freudenreichii* submetida a uma carência nutricional. Esses resultados, obtidos *in vitro* e em condições de laboratório, devem ser complementadas e validadas, estudando o comportamento de *P. freudenreichii* CIRM-BIA138 na matriz de queijo. Esta linhagem foi isolada à partir do queijo Emmental. A matriz de queijo é complexa, tem propriedades físico-químicas

específicas e as bactérias são imóveis, então, seria interessante estudar o comportamento de CIRM-BIA138 neste contexto. Isso iria avaliar, em primeiro lugar, a sua sobrevivência a longo prazo na matriz de queijo e, por outro, verificar se os mecanismos que se suspeita importante na transição para a fase estacionária tardia também estão envolvidos na sobrevivência a longo prazo em queijo.

**(VII) Ativação e regulação de vias metabólicas em *P. freudenreichii* CIRM-BIA138:**

Após o consumo total de lactato, o seu substrato preferencial, *P. freudenreichii* CIRM-BIA138 ativa várias outras vias metabólicas para garantir a sua sobrevivência. Seria interessante entender como estas vias alternativas são ativadas. Todas elas estão ativas simultaneamente ou são induzidas somente quando o lactato está esgotado? Neste caso, eles são induzidos por concomitante ou sequencial? Seria interessante, neste caso, identificar os estímulos e mecanismos reguladores. Uma opção é a construção de mutantes de genes envolvidos e induzidos no início de cada um dos processos metabólicos específicos. Por exemplo, para ativar o metabolismo de inositol, poderíamos tentar determinar se *P. freudenreichii* reconhece especificamente, a presença de inositol no meio e, em caso afirmativo, como esta bactéria distingue o estímulo no meio e reorienta o seu metabolismo. Isto poderia envolver a construção de genes mutantes com *iolT1* (*Major myo-inositol transporter iolT1*), *iolT2* (*Myo-inositol transporter iolT2*) e *iolT3* (*Myo-inositol transporter iolT3*). O acompanhamento por RT-qPCR, de outros genes envolvidos no metabolismo de inositol, e dosagens bioquímicas, da concentração de inositol nos sobrenadantes de cultura, podem ser realizadas nestes mutantes e podem ser comparadas a linhagem selvagem.

## Perspectives

Les résultats obtenus au cours de ce travail de thèse nous ont permis de mieux cerner le phénomène de survie à long terme chez *P. freudenreichii*. Ils ouvrent aussi plusieurs perspectives. D'une part, les données générées peuvent être complétées et exploitées pour améliorer notre connaissance sur la génomique de *P. freudenreichii* (points 1 et 2 ci-dessous). D'autre part, beaucoup reste encore à faire concernant l'analyse de la survie à long terme et plusieurs pistes pourraient permettre de compléter ce travail (points 3 à 7).

### (I) Achever la séquence du génome de *P. freudenreichii* CIRM-BIA138 :

En première approche, on peut considérer qu'un génome complet (c'est-à-dire, sans gaps) n'est pas indispensable si l'on peut en extraire les informations nécessaires aux analyses transcriptomique et métabolomique de cette bactérie. Cependant, un génome entièrement assemblé permettrait une meilleure annotation et donnerait ainsi une meilleure qualité d'informations génétiques sur l'organisme.

Pour obtenir le génome complet de *P. freudenreichii* CIRM-BIA138, une option serait d'effectuer une cartographie optique du génome de cette souche. Ceci permettrait de pallier les difficultés rencontrées dans l'étape d'assemblage de ce génome, dues à la présence d'une grande quantité de transposases. La cartographie optique donne en effet un aperçu général du génome, aide à la réorganisation de contigs et à la fermeture des gaps. En outre, il est possible de combiner différents programmes d'assemblage pour améliorer l'ensemble du processus. Par ailleurs, comme cela a déjà été démontré dans de précédentes études (Garber *et al.* 2011), et pour compléter les précédentes approches, il est possible d'utiliser les données transcriptomiques dérivées du RNA-Seq pour aligner les transcrits sur le génome et pour fermer les gaps. Les données de RNA-Seq peuvent également aider au processus d'annotation du génome, en utilisant des programmes tels que ReadXplorer qui permettent l'alignement des données d'expression de chaque gène sur le génome pour en affiner l'annotation (Hilker *et al.* 2014).

### (II) L'identification de nouveaux éléments dans le génome de *P. freudenreichii* CIRM-BIA 138 :

Les données de RNA-Seq permettent la découverte de nouveaux transcrits qui n'ont pas été détectés et annotés sur les précédentes versions du génome. Une analyse dédiée permettrait également d'identifier de petits ARN non codants chez *P.*



*freudenreichii* CIRM-BIA138. La technique de RNA-Seq s'est révélée efficace pour la détection de petits ARNs chez *Pseudomonas aeruginosa* (Gómez-Lozano *et al.* 2012), ainsi que pour la découverte de nouveaux transcrits dans plusieurs espèces bactériennes (McClure *et al.* 2013). Des programmes bioinformatiques, comme Rockhopper, grâce à une interface graphique conviviale, permettent à l'utilisateur d'y entrer des génomes de référence et de modifier des paramètres, tels que : les lectures *paired-end* ou *mate-pair* ; l'alignement des lectures de plusieurs réplicats ; le nombre des réplicats et de conditions expérimentales différentes utilisés pour les analyses de RNA-Seq (McClure *et al.* 2013).

### **(III) Compléter les données de RNA-Seq sur la phase stationnaire tardive de *P. freudenreichii* CIRM-BIA138 :**

Au cours de l'analyse RNA-Seq de la souche *P. freudenreichii* CIRM-BIA138, la quantité d'ARN extraite des échantillons liés à la phase de survie à long terme a été trop faible pour permettre un séquençage de qualité. De plus, les échantillons de la phase stationnaire tardive ont été contaminés par de l'ADN génomique. Ceci a fortement limité la portée des résultats obtenus. Il serait donc intéressant d'obtenir une plus grande quantité d'ARN et d'en améliorer la qualité, en augmentant le volume de culture bactérienne échantillonné (par ex. 25 mL) en vue d'améliorer, par la suite, les étapes d'extraction et de déplétion des ARNr en veillant à éliminer l'ADN génomique par des traitements à la DNase.

### **(IV) Le mécanisme de dormance dans l'espèce *P. freudenreichii* :**

De nombreuses bactéries à Gram positif et à Gram négatif entrent en dormance dans la phase de survie à long terme. Comme nous l'avons vu, parmi les 8 souches de *P. freudenreichii* testées, toutes survivent bien et leurs populations se maintiennent à des taux élevés au cours de la période de carence nutritionnelle. Nous avons cependant observé que la souche CIRM-BIA138 présentait un niveau de survie supérieur aux autres souches. Cette souche a fait l'objet d'une analyse plus poussée (démonstration de l'état VBNC et RNA-Seq). Dans ce contexte, il serait intéressant de compléter les résultats obtenus dans cette thèse, en explorant la capacité de survie à long terme et l'état VBNC d'autres souches. Ceci permettrait peut-être d'identifier des souches de *P. freudenreichii* présentant une faible survie à long terme puis d'identifier par génomique comparative les gènes impliqués (par exemple, par leur absence ou leur

pseudogénération) dans cette moindre survie. De même, d'autres stress (par exemple, thermique, acide, ou osmotique) pourraient être appliqués aux cultures en état de survie à long terme et carence nutritionnelle. Ceci permettrait d'identifier des phénomènes d'adaptation et de résistance croisées qui pourraient nous éclairer sur les mécanismes mis en œuvre et mieux expliquer le comportement de l'espèce en conditions de fabrication fromagère.

En outre, cette étude pourrait être complétée par la construction de mutants visant à interrompre les gènes impliqués dans le déclenchement de la dormance chez *P. freudenreichii*. Des travaux de ce type (construction de mutants) ont déjà été menés pour montrer l'implication du gène *clpX* chez *S. typhimurium* (Kusumoto, Miyashita and Kawamoto 2013) et *rpoS* chez *E. coli* (Boaretti *et al.* 2003) dans la survie à long terme chez ces espèces.

#### **(V) La réanimation dans l'espèce *P. freudenreichii* :**

Au sortir de l'état VBNC, de nombreuses espèces bactériennes passent par un état appelé réanimation leur permettant de reprendre une croissance normale. La réanimation (ou ressuscitation) nécessite parfois des conditions de croissance particulières. De même, certains gènes semblent directement impliqués dans ce phénomène. Ainsi, chez *M. luteus*, il a été montré que les cellules sont réanimées après la production du facteur Rpf (*Ressuscitation Promoting Factor*) (Greenblatt *et al.* 2004a). L'investigation de cet état de ressuscitation dans les cellules de *P. freudenreichii* serait d'un grand intérêt pour compléter ce travail. Cependant, il est nécessaire d'effectuer des expériences supplémentaires prouvant que les cellules sont réanimées après l'état VBNC. Cela requiert d'identifier les conditions de culture qui déclenchent la sortie de l'état VBNC ainsi que la reprise d'une croissance normale.

#### **(VI) La survie à long terme chez *P. freudenreichii* CIRM-BIA138 en matrice fromagère :**

Nos résultats permettent une meilleure compréhension des mécanismes moléculaires impliqués dans la survie à long terme de *P. freudenreichii* soumis à une carence nutritionnelle. Ces résultats, obtenus *in vitro* et en conditions de laboratoire, devraient être complétés et validés en étudiant le comportement de *P. freudenreichii* CIRM-BIA138 en matrice fromagère. Cette souche a été isolée à partir du fromage Emmental. La matrice fromagère est complexe, a des propriétés physico-chimiques

particulières et les bactéries y sont immobilisées, il serait donc intéressant d'étudier le comportement de CIRM-BIA138 dans ce contexte. Ceci permettrait d'évaluer, d'une part, sa survie à long terme en matrice fromagère et, d'autre part, de vérifier si les mécanismes que l'on suspecte importants dans la transition vers la phase stationnaire tardive sont aussi impliqués dans la survie à long terme en fromage.

**(VII) L'activation et la régulation des voies métaboliques chez *P. freudenreichii* CIRM-BIA138 :**

Après consommation totale du lactate, son substrat préférentiel, *P. freudenreichii* CIRM-BIA138 active plusieurs autres voies métaboliques pour assurer sa survie. Il serait intéressant de comprendre comment ces voies métaboliques alternatives sont activées. Sont-elles toutes actives en même temps ou sont-elles induites seulement lorsque le lactate est épuisé ? Dans ce cas, sont-elles induites de façon concomitante ou séquentielle ? Il serait intéressant dans ce cas d'en identifier les stimuli et mécanismes régulateurs. Une option consisterait à construire des mutants des gènes impliqués et induits au début de chaque processus métabolique spécifique. Par exemple, pour l'activation du métabolisme de l'inositol, nous pourrions essayer de déterminer si *P. freudenreichii* arrive à sentir spécifiquement la présence d'inositol dans le milieu et, si oui, comment cette bactérie reconnaît ce stimulus dans le milieu et réoriente son métabolisme. Ceci pourrait passer par la construction de mutants avec les gènes *iolT1* (*Major myo-inositol transporter iolT1*), *iolT2* (*Myo-inositol transporter iolT2*) et *iolT3* (*Myo-inositol transporter iolT3*). Le suivi, par RT-qPCR, des autres gènes impliqués dans le métabolisme de l'inositol et, par dosage biochimique, de la concentration d'inositol dans les surnageants de culture pourra ensuite être mené sur ces mutants et comparé à la souche sauvage.

## **Referências/ Références**

## Referências/ Références

- Abeijon Mukdsi MC, Falentin H, Maillard M-B *et al.* The Secreted Esterase of *Propionibacterium freudenreichii* Has a Major Role in Cheese Lipolysis. *Appl Environ Microbiol* 2014;**80**:751–6.
- Amit Singh AKG. Glycogenomics of *Mycobacterium tuberculosis*. *Mycobact Dis* 2014;**04**, DOI: 10.4172/2161-1068.1000175.
- Ammar EM, Jin Y, Wang Z *et al.* Metabolic engineering of *Propionibacterium freudenreichii*: effect of expressing phosphoenolpyruvate carboxylase on propionic acid production. *Appl Microbiol Biotechnol* 2014;**98**:7761–72.
- Anastasiou R, Leverrier P, Krestas I *et al.* Changes in protein synthesis during thermal adaptation of *Propionibacterium freudenreichii* subsp. *shermanii*. *Int J Food Microbiol* 2006;**108**:301–14.
- Arentz G, Weiland F, Oehler MK *et al.* State of the art of 2D DIGE. *Proteomics Clin Appl* 2014, DOI: 10.1002/prca.201400119.
- Backofen R, Amman F, Costa F *et al.* Bioinformatics of prokaryotic RNAs. *RNA Biol* 2014;**11**:470–83.
- Balows A ed. *The Prokaryotes: A Handbook on the Biology of Bacteria: Ecophysiology, Isolation, Identification, Applications*. New York: Springer-Verlag, 1992.
- Berdagué J, Grappin R, Delacroix-Buchet A *et al.* Caractérisation de l’Emmental “Grand-Cru” français. I. Composition physico-chimique. *Le Lait* 1990;**70**:1–14.
- Bernstein CN. Antibiotics, Probiotics and Prebiotics in IBD. In: Lewis JD, Ruemmele FM, Wu GD (eds.). *Nestlé Nutrition Institute Workshop Series*. Vol 79. Basel: S. KARGER AG, 2014, 83–100.
- Bharati BK, Sharma IM, Kasetty S *et al.* A full-length bifunctional protein involved in c-di-GMP turnover is required for long-term survival under nutrient starvation in *Mycobacterium smegmatis*. *Microbiology* 2012;**158**:1415–27.
- Boaretti M, Lleò MM, Bonato B *et al.* Involvement of *rpoS* in the survival of *Escherichia coli* in the viable but non-culturable state. *Environ Microbiol* 2003;**5**:986–96.
- Boehm M, Slack F. A Developmental Timing MicroRNA and Its Target Regulate Life Span in *C. elegans*. *Science* 2005;**310**:1954–7.
- Bourassa L, Camilli A. Glycogen contributes to the environmental persistence and transmission of *Vibrio cholerae*. *Mol Microbiol* 2009;**72**:124–38.

- Bron PA, van Baarlen P, Kleerebezem M. Emerging molecular insights into the interaction between probiotics and the host intestinal mucosa. *Nat Rev Microbiol* 2012;**10**:66–78.
- Buchanan RE. Life Phases in a Bacterial Culture. *J Infect Dis* 1918;**23**:109–25.
- Burand JP, Hunter WB. RNAi: Future in insect management. *J Invertebr Pathol* 2013;**112**, **Supplement 1**:S68–74.
- Carvalho M, Oehmen A, Carvalho G *et al.* Survival strategies of polyphosphate accumulating organisms and glycogen accumulating organisms under conditions of low organic loading. *Bioresour Technol* 2014;**172**:290–6.
- de Carvalho AF, Gautier M, Grimont F. Identification of dairy Propionibacterium species by rRNA gene restriction patterns. *Res Microbiol* 1994;**145**:667–76.
- Champomier Vergès M-C, Zúñiga M, Morel-Deville F *et al.* Relationships between arginine degradation, pH and survival in *Lactobacillus sakei*. *FEMS Microbiol Lett* 1999;**180**:297–304.
- Chien A-C, Hill NS, Levin PA. Cell Size Control in Bacteria. *Curr Biol* 2012;**22**:R340–9.
- Cousin FJ, Jouan-Lanhouet S, Dimanche-Boitrel M-T *et al.* Milk fermented by *Propionibacterium freudenreichii* induces apoptosis of HGT-1 human gastric cancer cells. *PLoS One* 2012;**7**:e31892.
- Cousin FJ, Mater DDG, Foligné B *et al.* Dairy propionibacteria as human probiotics: A review of recent evidence. *Dairy Sci Technol* 2010;**91**:1–26.
- Creedy JP, Conway T. Quantitative bacterial transcriptomics with RNA-seq. *Curr Opin Microbiol* 2015;**23**:133–40.
- Crow VL. Metabolism of Aspartate by *Propionibacterium freudenreichii* subsp. *shermanii*: Effect on Lactate Fermentation. *Appl Environ Microbiol* 1986;**52**:359–65.
- Cummins C, Johnson J. The genus *Propionibacterium*. *The prokaryotes*.
- Dalmaso M, Aubert J, Briard-Bion V *et al.* A Temporal -omic Study of *Propionibacterium freudenreichii* CIRM-BIA1T Adaptation Strategies in Conditions Mimicking Cheese Ripening in the Cold. Gilbert JA (ed.). *PLoS ONE* 2012a;**7**:e29083.
- Dalmaso M, Aubert J, Briard-Bion V *et al.* Correction: A Temporal -omic Study of *Propionibacterium freudenreichii* CIRM-BIA1T Adaptation Strategies in Conditions Mimicking Cheese Ripening in the Cold. *PLoS ONE* 2012b;**7**:10.1371/annotation/e0ff065d – a52d – 44f2–8727 – 328393ed60b6.
- Dalmaso M, Nicolas P, Falentin H *et al.* Multilocus sequence typing of *Propionibacterium freudenreichii*. *Int J Food Microbiol* 2011;**145**:113–20.

- Davis C. Enumeration of probiotic strains: Review of culture-dependent and alternative techniques to quantify viable bacteria. *J Microbiol Methods* 2014;**103**:9–17.
- Del Chierico F, Ancora M, Marcacci M *et al.* Choice of next-generation sequencing pipelines. *Bacterial Pangenomics*. Springer, 2015, 31–47.
- Dewi Puspita I, Uehara M, Katayama T *et al.* Resuscitation Promoting Factor (Rpf) from *Tomitella biformata* AHU 1821T Promotes Growth and Resuscitates Non-Dividing Cells. *Microbes Environ* 2013;**28**:58–64.
- Dherbécourt J, Falentin H, Canaan S *et al.* A genomic search approach to identify esterases in *Propionibacterium freudenreichii* involved in the formation of flavour in Emmental cheese. *Microb Cell Factories* 2008;**7**:16.
- Diaper JP, Edwards C. Survival of *Staphylococcus aureus* in lakewater monitored by flow cytometry. *Microbiol Read Engl* 1994;**140** ( Pt 1):35–42.
- EFSA experts. The maintenance of the list of QPS microorganisms intentionally added to food or feed - Scientific Opinion of the Panel on Biological Hazards. *EFSA Journal*. 2008.
- Falentin H, Deutsch S-M, Jan G *et al.* The Complete Genome of *Propionibacterium freudenreichii* CIRM-BIA1T, a Hardy Actinobacterium with Food and Probiotic Applications. Ahmed N (ed.). *PLoS ONE* 2010a;**5**:e11748.
- Falentin H, Deutsch S-M, Jan G *et al.* The Complete Genome of *Propionibacterium freudenreichii* CIRM-BIA1T, a Hardy Actinobacterium with Food and Probiotic Applications. *PLoS ONE* 2010b;**5**:e11748.
- Falentin H, Postollec F, Parayre S *et al.* Specific metabolic activity of ripening bacteria quantified by real-time reverse transcription PCR throughout Emmental cheese manufacture. *Int J Food Microbiol* 2010c;**144**:10–9.
- Fine AE, Bolin CA, Gardiner JC *et al.* A Study of the Persistence of *Mycobacterium bovis* in the Environment under Natural Weather Conditions in Michigan, USA, A Study of the Persistence of *Mycobacterium bovis* in the Environment under Natural Weather Conditions in Michigan, USA. *Vet Med Int Vet Med Int* 2011;**2011**, **2011**:e765430.
- Finkel SE. Long-term survival during stationary phase: evolution and the GASP phenotype. *Nat Rev Microbiol* 2006a;**4**:113–20.
- Finkel SE. Long-term survival during stationary phase: evolution and the GASP phenotype. *Nat Rev Microbiol* 2006b;**4**:113–20.
- Finkel SE. Long-term survival during stationary phase: evolution and the GASP phenotype. *Nat Rev Microbiol* 2006c;**4**:113–20.
- Florent. Vitamins. *Biotechnology*. 1986:119–58.

- Fouquier d'Herouel A, Wessner F, Halpern D *et al.* A simple and efficient method to search for selected primary transcripts: non-coding and antisense RNAs in the human pathogen *Enterococcus faecalis*. *Nucleic Acids Res* 2011;**39**:e46–e46.
- Garber M, Grabherr MG, Guttman M *et al.* Computational methods for transcriptome annotation and quantification using RNA-seq. *Nat Methods* 2011;**8**:469–77.
- García Hernández F, Pérez Mendoza JL. [Fermentation of agro-industrial wastes for obtaining metabolites of industrial interest]. *Bol Estud Méd Biológicos* 1983;**32**:177–95.
- Geisel N, Vilar JMG, Rubi JM. Optimal Resting-Growth Strategies of Microbial Populations in Fluctuating Environments. *PLoS ONE* 2011;**6**:e18622.
- Geuking MB, Köller Y, Rupp S *et al.* The interplay between the gut microbiota and the immune system. *Gut Microbes* 2014;**5**:411–8.
- Gómez-Lozano M, Marvig RL, Molin S *et al.* Genome-wide identification of novel small RNAs in *Pseudomonas aeruginosa*. *Environ Microbiol* 2012;**14**:2006–16.
- Greenblatt CL, Baum J, Klein BY *et al.* *Micrococcus luteus* - Survival in Amber. *Microb Ecol* 2004a;**48**:120–7.
- Greenblatt CL, Baum J, Klein BY *et al.* *Micrococcus luteus* - Survival in Amber. *Microb Ecol* 2004b;**48**:120–7.
- Haas BJ, Chin M, Nusbaum C *et al.* How deep is deep enough for RNA-Seq profiling of bacterial transcriptomes? *BMC Genomics* 2012;**13**:734.
- Hague A, Elder DJ, Hicks DJ *et al.* Apoptosis in colorectal tumour cells: induction by the short chain fatty acids butyrate, propionate and acetate and by the bile salt deoxycholate. *Int J Cancer J Int Cancer* 1995;**60**:400–6.
- Helmus RA, Liermann LJ, Brantley SL *et al.* Growth advantage in stationary-phase (GASP) phenotype in long-term survival strains of *Geobacter sulfurreducens*. *FEMS Microbiol Ecol* 2012;**79**:218–28.
- Hervé C, Fondrevez M, Chéron A *et al.* Transcarboxylase mRNA: a marker which evidences *P. freudenreichii* survival and metabolic activity during its transit in the human gut. *Int J Food Microbiol* 2007;**113**:303–14.
- Hilker R, Stadermann KB, Doppmeier D *et al.* ReadXplorer—visualization and analysis of mapped sequences. *Bioinformatics* 2014:btu205.
- Hooper LV, Macpherson AJ. Immune adaptations that maintain homeostasis with the intestinal microbiota. *Nat Rev Immunol* 2010;**10**:159–69.
- Huang Y, Adams MC. In vitro assessment of the upper gastrointestinal tolerance of potential probiotic dairy propionibacteria. *Int J Food Microbiol* 2004;**91**:253–60.



- Isawa K, HOJO K, YODAN *et al.* Isolation and Identification of a New Bifidogenic Growth Stimulator Produced by *Propionibacterium freudenreichii* ET-3. *Biosci Biotechnol Biochem* 2002;**66**:679–81.
- Jan G, Belzacq A-S, Haouzi D *et al.* Propionibacteria induce apoptosis of colorectal carcinoma cells via short-chain fatty acids acting on mitochondria. *Cell Death Differ* 2002;**9**:179–88.
- Jan G, Leverrier P, Pichereau V *et al.* Changes in Protein Synthesis and Morphology during Acid Adaptation of *Propionibacterium freudenreichii*. *Appl Environ Microbiol* 2001;**67**:2029–36.
- Jan G, Rouault A, Maubois J-L. Acid stress susceptibility and acid adaptation of *Propionibacterium freudenreichii* subsp. *shermanii*. *Le Lait* 2000;**80**:325–36.
- Johnson DS, Mortazavi A, Myers RM *et al.* Genome-wide mapping of in vivo protein-DNA interactions. *Science* 2007;**316**:1497–502.
- Kato M, Chen X, Inukai S *et al.* Age-associated changes in expression of small, noncoding RNAs, including microRNAs, in *C. elegans*. *Rna* 2011;**17**:1804–20.
- Koussémon M, Combet-Blanc Y, Patel BK *et al.* *Propionibacterium microaerophilum* sp. nov., a microaerophilic bacterium isolated from olive mill wastewater. *Int J Syst Evol Microbiol* 2001;**51**:1373–82.
- Kröger C, Dillon SC, Cameron ADS *et al.* The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Typhimurium. *Proc Natl Acad Sci* 2012;**109**:E1277–86.
- Kusano K, Yamada H, Niwa M *et al.* *Propionibacterium cyclohexanicum* sp. nov., a new acid-tolerant omega-cyclohexyl fatty acid-containing propionibacterium isolated from spoiled orange juice. *Int J Syst Bacteriol* 1997;**47**:825–31.
- Kusumoto A, Miyashita M, Kawamoto K. Deletion in the C-terminal domain of ClpX delayed entry of *Salmonella enterica* into a viable but non-culturable state. *Res Microbiol* 2013;**164**:335–41.
- Lan A, Bruneau A, Bensaada M *et al.* Increased induction of apoptosis by *Propionibacterium freudenreichii* TL133 in colonic mucosal crypts of human microbiota-associated rats treated with 1,2-dimethylhydrazine. *Br J Nutr* 2008;**100**:1251–9.
- Lan A, Bruneau A, Philippe C *et al.* Survival and metabolic activity of selected strains of *Propionibacterium freudenreichii* in the gastrointestinal tract of human microbiota-associated rats. *Br J Nutr* 2007;**97**:714.
- Langsrud and Reinbold. Flavor development and microbiology of Swiss cheese- A review. III. Ripening and flavor production. *J Milk Food Technol* 1973;**36**:593–609.

Le Boucher C, Courant F, Jeanson S *et al.* First mass spectrometry metabolic fingerprinting of bacterial metabolism in a model cheese. *Food Chem* 2013;**141**:1032–40.

Le Marechal C, Peton V, Ple C *et al.* Surface proteins of *Propionibacterium freudenreichii* are involved in its anti-inflammatory properties. *J Proteomics* 2015;**113**:447–61.

Le Maréchal C, Peton V, Plé C *et al.* Surface proteins of *Propionibacterium freudenreichii* are involved in its anti-inflammatory properties. *J Proteomics* 2015;**113**:447–61.

Lemunier M, Francou C, Rousseaux S *et al.* Long-term survival of pathogenic and sanitation indicator bacteria in experimental biowaste composts. *Appl Environ Microbiol* 2005;**71**:5779–86.

Leverrier P, Dimova D, Pichereau V *et al.* Susceptibility and adaptive response to bile salts in *Propionibacterium freudenreichii*: physiological and proteomic analysis. *Appl Environ Microbiol* 2003a;**69**:3809–18.

Leverrier P, Dimova D, Pichereau V *et al.* Susceptibility and Adaptive Response to Bile Salts in *Propionibacterium freudenreichii*: Physiological and Proteomic Analysis. *Appl Environ Microbiol* 2003b;**69**:3809–18.

Lilly DM, Stillwell RH. PROBIOTICS: GROWTH-PROMOTING FACTORS PRODUCED BY MICROORGANISMS. *Science* 1965;**147**:747–8.

Loux L, Mariadassou M, Almeida S *et al.* Mutations and genomic islands explain the strain dependency of sugar utilization among 21 strains of *Propionibacterium freudenreichii*. *BMC Genomics* 2015a.

Loux V, Mariadassou M, Almeida S *et al.* Mutations and genomic islands can explain the strain dependency of sugar utilization in 21 strains of *Propionibacterium freudenreichii*. *BMC Genomics* 2015b;**16**:296.

Luerce T, Gomes-Santos A, Rocha C *et al.* Anti-inflammatory effects of *Lactococcus lactis* NCDO 2118 during the remission period of chemically induced colitis. *Gut Pathog* 2014;**6**:33.

Luijk NV, Stierli MP, Schwenninger SM *et al.* Genetics and molecular biology of propionibacteria. *Le Lait* 2002;**82**:14.

Malik AC, Reinbold GW, Vedamuthu ER. Evaluation of the taxonomy of the *Propionibacterium*. *Canadian Journal of Microbiology*. 1968:1185–91.

Mantere-Alhonen S. Die Propionibakterien der Molkereiindustrie als Darmkanalmikroben. *Meijeritiet Aikak* 1982;**40**.

- Mantere-Alhonen S. Propionibacteria used as probiotics - A review. *Le Lait* 1995;**75**:447–52.
- Martin JA, Wang Z. Next-generation transcriptome assembly. *Nat Rev Genet* 2011;**12**:671–82.
- Marwaha SS, Sethi RP, Kennedy JF. Influence of 5,6-dimethylbenzimidazole (DMB) on vitamin B12 biosynthesis by strains of Propionibacterium. *Enzyme Microb Technol* 1983;**5**:361–4.
- McClure R, Balasubramanian D, Sun Y *et al.* Computational analysis of bacterial RNA-Seq data. *Nucleic Acids Res* 2013:gkt444.
- Meile L, Dasen G, Miescher S *et al.* Classification of propionic acid bacteria and approaches to applied genetics. *Le Lait* 1999;**79**:71–8.
- Meurice G, Jacob D, Deborde C *et al.* Whole genome sequencing project of a dairy Propionibacterium freudenreichii subsp. shermanii genome: progress and first bioinformatic analysis. *Le Lait* 2004;**84**:15–24.
- Mogensen G, SALMINEN;S, O'BRIEN;J *et al.* *Inventory of Microorganisms with a Documented History of Use in Food*. Brussels, BELGIQUE: International Dairy Federation, 2002.
- Moore W, Cato E. VALIDITY OF PROPIONIBACTERIUM ACNES (GILCHRIST) DOUGLAS AND GUNTER COMB. NOV. *Journal of Bacteriology*. 1963:870–4.
- Mulyukin AL, Soina VS, Demkina EV *et al.* Formation of resting cells by non-spore-forming microorganisms as a strategy of long-term survival in the environment. In: Hoover RB, Rozanov AY, Lipps JH (eds.). 2003, 208–18.
- Murooka Y, Piao Y, Kiatpapan P *et al.* Production of tetrapyrrole compounds and vitamin B12 using genetically engineering of Propionibacterium freudenreichii. An overview. *Le Lait* 2005;**85**:14.
- Nelson DL, Nelson DL, Lehninger AL *et al.* *Lehninger Principles of Biochemistry*. New York: W.H. Freeman, 2008.
- Okada Y, Tsuzuki Y, Miyazaki J *et al.* Propionibacterium freudenreichii component 1,4-dihydroxy-2-naphthoic acid (DHNA) attenuates dextran sodium sulphate induced colitis by modulation of bacterial flora and lymphocyte homing. *Gut* 2006;**55**:681–8.
- Okada Y, Tsuzuki Y, Narimatsu K *et al.* 1,4-Dihydroxy-2-naphthoic acid from Propionibacterium freudenreichii reduces inflammation in interleukin-10-deficient mice with colitis by suppressing macrophage-derived proinflammatory cytokines. *J Leukoc Biol* 2013;**94**:473–80.

- Olszewska M, Staniewski B, Łaniewska-Trokenheim Ł. Cell Viability of Bifidobacterium lactis Strain in Long-Term Storage Butter Assessed with the Plate Count and Fluorescence Techniques. *Czech J Food Sci* 2012;**30**:421–8.
- Orla-Jensen. *Die Hauptlinien des natürlichen Bakteriensystems nebst einer Uebersicht der Gärungsphänomene.*, 1909.
- Ouwehand AC, Kirjavainen PV, Grönlund M-M *et al.* Adhesion of probiotic microorganisms to intestinal mucus. *Int Dairy J* 1999;**9**:623–30.
- Ouwehand AC, Suomalainen T, Tölkö S *et al.* In vitro adhesion of propionic acid bacteria to human intestinal mucus. *Le Lait* 2002;**82**:8.
- Ouwehand AC, Tiihonen K, Saarinen M *et al.* Influence of a combination of Lactobacillus acidophilus NCFM and lactitol on healthy elderly: intestinal and immune parameters. *Br J Nutr* 2009;**101**:367–75.
- Papenfert K, Vogel J. Regulatory RNA in Bacterial Pathogens. *Cell Host Microbe* 2010;**8**:116–27.
- Parayre S, Falentin H, Madec M-N *et al.* Easy DNA extraction method and optimisation of PCR-Temporal Temperature Gel Electrophoresis to identify the predominant high and low GC-content bacteria from dairy products. *J Microbiol Methods* 2007;**69**:431–41.
- Parizzi LP, Grassi MCB, Llerena LA *et al.* The genome sequence of Propionibacterium acidipropionici provides insights into its biotechnological and industrial potential. *BMC Genomics* 2012;**13**:562.
- Park PJ. ChIP-seq: advantages and challenges of a maturing technology. *Nat Rev Genet* 2009;**10**:669–80.
- Parvez S, Malik KA, Ah Kang S *et al.* Probiotics and their fermented food products are beneficial for health. *J Appl Microbiol* 2006;**100**:1171–85.
- Pérez Chaia A, Zarate G. Dairy propionibacteria from milk or cheese diets remain viable and enhance propionic acid production in the mouse cecum. *Le Lait* 2005;**85**:85–98.
- Pérez Chaia A, Zarate G, Oliver G. The probiotic properties of propionibacteria. *Le Lait* 1999;**79**:175–85.
- Piao Y, Yamashita M, Kawaraichi N *et al.* Production of vitamin B12 in genetically engineered Propionibacterium freudenreichii. *J Biosci Bioeng* 2004;**98**:167–73.
- Picard C, Fioramonti J, Francois A *et al.* Review article: bifidobacteria as probiotic agents – physiological effects and clinical benefits. *Aliment Pharmacol Ther* 2005;**22**:495–512.

- Pinto AC, Sá PHCG de, Ramos RTJ *et al.* Differential transcriptional profile of *Corynebacterium pseudotuberculosis* in response to abiotic stresses. *BMC Genomics* 2014;**15**:14.
- Poonam, Pophaly SD, Tomar SK *et al.* Multifaceted attributes of dairy propionibacteria: a review. *World J Microbiol Biotechnol* 2012;**28**:3081–95.
- Primm TP, Andersen SJ, Mizrahi V *et al.* The Stringent Response of *Mycobacterium tuberculosis* Is Required for Long-Term Survival. *J Bacteriol* 2000;**182**:4889–98.
- Quigley ME, Hudson GJ, Englyst HN. Determination of resistant short-chain carbohydrates (non-digestible oligosaccharides) using gas–liquid chromatography. *Food Chem* 1999;**65**:381–90.
- Ramamurthy T, Ghosh A, Pazhani GP *et al.* Current Perspectives on Viable but Non-Culturable (VBNC) Pathogenic Bacteria. *Front Public Health* 2014;**2**, DOI: 10.3389/fpubh.2014.00103.
- Rao NN, Gómez-García MR, Kornberg A. Inorganic Polyphosphate: Essential for Growth and Survival. *Annu Rev Biochem* 2009;**78**:605–47.
- Ratcliff WC, Hoverman M, Travisano M *et al.* Disentangling direct and indirect fitness effects of microbial dormancy. *Am Nat* 2013;**182**:147–56.
- Rifat D, Bishai WR, Karakousis PC. Phosphate depletion: a novel trigger for *Mycobacterium tuberculosis* persistence. *J Infect Dis* 2009;**200**:1126–35.
- Rittershaus ESC, Baek S-H, Sasseti CM. The Normalcy of Dormancy: Common Themes in Microbial Quiescence. *Cell Host Microbe* 2013a;**13**:643–51.
- Rittershaus ESC, Baek S, Sasseti CM. The Normalcy of Dormancy. *Cell Host Microbe* 2013b;**13**:643–51.
- Robertson G, Hirst M, Bainbridge M *et al.* Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat Methods* 2007;**4**:651–7.
- Rozzak DB, Colwell RR. Survival strategies of bacteria in the natural environment. *Microbiol Rev* 1987;**51**:365–79.
- Santos Rocha C, Lakhdari O, Blottière HM *et al.* Anti-inflammatory properties of dairy lactobacilli. *Inflamm Bowel Dis* 2012;**18**:657–66.
- Saraoui T, Parayre S, Guernec G *et al.* A unique in vivo experimental approach reveals metabolic adaptation of the probiotic *Propionibacterium freudenreichii* to the colon environment. *BMC Genomics* 2013a;**14**:911.

- Saraoui T, Parayre S, Guernec G *et al.* A unique in vivo experimental approach reveals metabolic adaptation of the probiotic *Propionibacterium freudenreichii* to the colon environment. *BMC Genomics* 2013b;**14**:911.
- Sartor RB. Microbial Influences in Inflammatory Bowel Diseases. *Gastroenterology* 2008;**134**:577–94.
- Sathish JG, Sethu S, Bielsky M-C *et al.* Challenges and approaches for the development of safer immunomodulatory biologics. *Nat Rev Drug Discov* 2013;**12**:306–24.
- Savic DJ, McShan WM. Long-term survival of *Streptococcus pyogenes* in rich media is pH-dependent. *Microbiology* 2012;**158**:1428–36.
- Seyednasrollah F, Laiho A, Elo LL. Comparison of software packages for detecting differential expression in RNA-seq studies. *Brief Bioinform* 2013, DOI: 10.1093/bib/bbt086.
- Sharma CM, Hoffmann S, Darfeuille F *et al.* The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* 2010;**464**:250–5.
- S. Kaprelyants A, V. Mukamolova G, Ruggiero A *et al.* Resuscitation-promoting Factors (Rpf): In Search of Inhibitors. *Protein Pept Lett* 2012;**19**:1026–34.
- Slavchev G, Michailova L, Markova N. Stress-induced L-forms of *Mycobacterium bovis*: a challenge to survivability. *New Microbiol* 2013;**36**:157–66.
- Stowers CC, Cox BM, Rodriguez BA. Development of an industrializable fermentation process for propionic acid production. *J Ind Microbiol Biotechnol* 2014;**41**:837–52.
- Suomalainen T, Lagström H, Mättö J *et al.* Influence of whey-based fruit juice containing *Lactobacillus rhamnosus* on intestinal well-being and humoral immune response in healthy adults. *LWT - Food Sci Technol* 2006;**39**:788–95.
- Thierry A, Deutsch S-M, Falentin H *et al.* New insights into physiology and metabolism of *Propionibacterium freudenreichii*. *Int J Food Microbiol* 2011;**149**:19–27.
- Tjaden B. De novo assembly of bacterial transcriptomes from RNA-seq data. *Genome Biol* 2015;**16**:1.
- Vallenet D, Nordmann P, Barbe V *et al.* Comparative Analysis of Acinetobacters: Three Genomes for Three Lifestyles. Hall N (ed.). *PLoS ONE* 2008;**3**:e1805.
- Vandamme E. Biotechnology of vitamins, pigments and growth factors. *Elsevier*. 1989:257–84.
- Vandamme EJ. Production of vitamins, coenzymes and related biochemicals by biotechnological processes. *J Chem Technol Biotechnol* 1992;**53**:313–27.

- Vangay P, Ward T, Gerber JS *et al.* Antibiotics, Pediatric Dysbiosis, and Disease. *Cell Host Microbe* 2015;**17**:553–64.
- Van Wyk J, Witthuhn RC, Britz TJ. Optimisation of vitamin B12 and folate production by *Propionibacterium freudenreichii* strains in kefir. *Int Dairy J* 2011;**21**:69–74.
- Velasquez-Manoff M. Gut Microbiome: The Peacekeepers. *Nature* 2015;**518**:S3–11.
- Vergara-Irigaray M, Fookes MC, Thomson NR *et al.* RNA-seq analysis of the influence of anaerobiosis and FNR on *Shigella flexneri*. *BMC Genomics* 2014;**15**:438.
- von Freudenreich and Orla-Jensen. Über die in emmentalerkäse stattfindene propionsäure-gärung. *Bakteriol* 1906;**17**:529–46.
- Vorobjeva LI. *Propionibacteria*. *Kluwer Acad Publ* 1999.
- de Vos WM, de Vos EAJ. Role of the intestinal microbiome in health and disease: from correlation to causation. *Nutr Rev* 2012;**70** **Suppl 1**:S45–56.
- Wang P, Zhang Z, Jiao Y *et al.* Improved propionic acid and 5,6-dimethylbenzimidazole control strategy for vitamin B12 fermentation by *Propionibacterium freudenreichii*. *J Biotechnol* 2015;**193**:123–9.
- Wang Z, Yang S-T. Propionic acid production in glycerol/glucose co-fermentation by *Propionibacterium freudenreichii* subsp. *shermanii*. *Bioresour Technol* 2013;**137**:116–23.
- Warminska-Radyko I, Laniewska-Moroz L, Babuchowski A. Possibilities for stimulation of *Bifidobacterium* growth by propionibacteria. *Le Lait* 2002;**82**:10.
- Watson SP, Clements MO, Foster SJ. Characterization of the starvation-survival response of *Staphylococcus aureus*. *J Bacteriol* 1998;**180**:1750–8.
- Wayne LG, Sohaskey CD. Nonreplicating persistence of mycobacterium tuberculosis. *Annu Rev Microbiol* 2001;**55**:139–63.
- Wen J, Deng X, Li Z *et al.* Transcriptomic Response of *Listeria monocytogenes* during the Transition to the Long-Term-Survival Phase. *Appl Environ Microbiol* 2011;**77**:5966–72.
- Wen J, Karthikeyan S, Hawkins J *et al.* *Listeria monocytogenes* responds to cell density as it transitions to the long-term-survival phase. *Int J Food Microbiol* 2013;**165**:326–31.
- Wheat WH, Casali AL, Thomas V *et al.* Long-term survival and virulence of *Mycobacterium leprae* in amoebal cysts. *PLoS Negl Trop Dis* 2014;**8**:e3405.
- Wilkinson JF. The problem of energy-storage compounds in bacteria. *Exp Cell Res* 1959;**7**, **Supplement**:111–30.

- Wood HG. Metabolic cycles in the fermentation by propionic acid bacteria. *Curr Top Cell Regul* 1981;**18**:255–87.
- Zarate G, Chaia AP, Oliver G. Some characteristics of practical relevance of the beta-galactosidase from potential probiotic strains of *Propionibacterium acidipropionici*. *Anaerobe* 2002;**8**:259–67.
- Zárate G, Morata De Ambrosini V, Perez Chaia A *et al.* Some factors affecting the adherence of probiotic *Propionibacterium acidipropionici* CRL 1198 to intestinal epithelial cells. *Can J Microbiol* 2002;**48**:449–57.
- Zinser ER, Kolter R. Mutations enhancing amino acid catabolism confer a growth advantage in stationary phase. *J Bacteriol* 1999;**181**:5800–7.
- Zinser ER, Kolter R. *Escherichia coli* evolution during stationary phase. *Res Microbiol* 2004;**155**:328–36.
- Zoephel J, Randau L. RNA-Seq analyses reveal CRISPR RNA processing and regulation patterns. *Biochem Soc Trans* 2013;**41**:1459–63.
- Ammar EM, Jin Y, Wang Z *et al.* Metabolic engineering of *Propionibacterium freudenreichii*: effect of expressing phosphoenolpyruvate carboxylase on propionic acid production. *Appl Microbiol Biotechnol* 2014;**98**:7761–72.
- Anastasiou R, Leverrier P, Krestas I *et al.* Changes in protein synthesis during thermal adaptation of *Propionibacterium freudenreichii* subsp. *shermanii*. *Int J Food Microbiol* 2006;**108**:301–14.
- Arentz G, Weiland F, Oehler MK *et al.* State of the art of 2D DIGE. *Proteomics Clin Appl* 2014, DOI: 10.1002/prca.201400119.
- Backofen R, Amman F, Costa F *et al.* Bioinformatics of prokaryotic RNAs. *RNA Biol* 2014;**11**:470–83.
- Balows A ed. *The Prokaryotes: A Handbook on the Biology of Bacteria: Ecophysiology, Isolation, Identification, Applications*. New York: Springer-Verlag, 1992.
- Barinov A, Loux V, Hammani A *et al.* Prediction of surface exposed proteins in *Streptococcus pyogenes*, with a potential application to other Gram-positive bacteria. *PROTEOMICS* 2009;**9**:61–73.
- Batista PJ, Chang HY. Long Noncoding RNAs: Cellular Address Codes in Development and Disease. *Cell* 2013;**152**:1298–307.
- Baum JA, Bogaert T, Clinton W *et al.* Control of coleopteran insect pests through RNA interference. *Nat Biotechnol* 2007;**25**:1322–6.
- Berdagué J, Grappin R, Delacroix-Buchet A *et al.* Caractérisation de l’Emmental “Grand-Cru” français. I. Composition physico-chimique. *Le Lait* 1990;**70**:1–14.



- Bernstein CN. Antibiotics, Probiotics and Prebiotics in IBD. In: Lewis JD, Ruemmele FM, Wu GD (eds.). *Nestlé Nutrition Institute Workshop Series*. Vol 79. Basel: S. KARGER AG, 2014, 83–100.
- den Besten G, van Eunen K, Groen AK *et al.* The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *J Lipid Res* 2013;**54**:2325–40.
- Bisanz JE, Macklaim JM, Gloor GB *et al.* Bacterial metatranscriptome analysis of a probiotic yogurt using an RNA-Seq approach. *Int Dairy J* 2014;**39**:284–92.
- Bischler T, Siew Tan H, Nieselt K *et al.* Differential RNA-seq (dRNA-seq) for annotation of transcriptional start sites and small RNAs in *Helicobacter pylori*. *Methods* 2015, DOI: 10.1016/j.ymeth.2015.06.012.
- Boaretti M, Lleò MM, Bonato B *et al.* Involvement of *rpoS* in the survival of *Escherichia coli* in the viable but non-culturable state. *Environ Microbiol* 2003;**5**:986–96.
- Boden D, Pusch O, Lee F *et al.* Human immunodeficiency virus type 1 escape from RNA interference. *J Virol* 2003;**77**:11531–5.
- Boehm M, Slack F. A Developmental Timing MicroRNA and Its Target Regulate Life Span in *C. elegans*. *Science* 2005;**310**:1954–7.
- Brendehaug J, Langsrud T. Amino Acid Metabolism in Propionibacteria: Resting Cells Experiments with Four Strains. *J Dairy Sci* 1985;**68**:281–9.
- Bron PA, van Baarlen P, Kleerebezem M. Emerging molecular insights into the interaction between probiotics and the host intestinal mucosa. *Nat Rev Microbiol* 2012;**10**:66–78.
- Burand JP, Hunter WB. RNAi: Future in insect management. *J Invertebr Pathol* 2013;**112**, Supplement 1:S68–74.
- Caboche S, Audebert C, Lemoine Y *et al.* Comparison of mapping algorithms used in high-throughput sequencing: application to Ion Torrent data. *BMC Genomics* 2014;**15**:264.
- de Carvalho AF, Gautier M, Grimont F. Identification of diary Propionibacterium species by rRNA gene restriction patterns. *Res Microbiol* 1994;**145**:667–76.
- Casals F, Idaghdour Y, Hussin J *et al.* Next-generation sequencing approaches for genetic mapping of complex diseases. *J Neuroimmunol* 2012;**248**:10–22.
- Chi SW, Zang JB, Mele A *et al.* Argonaute HITS-CLIP decodes microRNA–mRNA interaction maps. *Nature* 2009;**460**:479–86.
- Cho S, Cho Y, Lee S *et al.* Current Challenges in Bacterial Transcriptomics. *Genomics Inform* 2013;**11**:76.

- Chuang L-Y, Chang H-W, Tsai J-H *et al.* Features for computational operon prediction in prokaryotes. *Brief Funct Genomics* 2012;els024.
- Chu Y, Corey DR. RNA sequencing: platform selection, experimental design, and data interpretation. *Nucleic Acid Ther* 2012;**22**:271–4.
- Cousin FJ, Jouan-Lanhouet S, Dimanche-Boitrel M-T *et al.* Milk fermented by *Propionibacterium freudenreichii* induces apoptosis of HGT-1 human gastric cancer cells. *PLoS One* 2012;**7**:e31892.
- Cousin FJ, Mater DDG, Foligne B *et al.* Dairy propionibacteria as human probiotics: A review of recent evidence. *Dairy Sci Technol* 2010a, DOI: 10.1051/dst/2010032.
- Cousin FJ, Mater DDG, Foligné B *et al.* Dairy propionibacteria as human probiotics: A review of recent evidence. *Dairy Sci Technol* 2010b;**91**:1–26.
- Creecy JP, Conway T. Quantitative bacterial transcriptomics with RNA-seq. *Curr Opin Microbiol* 2015;**23**:133–40.
- Crow VL. Metabolism of Aspartate by *Propionibacterium freudenreichii* subsp. *shermanii*: Effect on Lactate Fermentation. *Appl Environ Microbiol* 1986;**52**:359–65.
- Cummins C, Johnson J. The genus *Propionibacterium*. *The prokaryotes*.
- Dai L, Gao X, Guo Y *et al.* Bioinformatics clouds for big data manipulation. *Biol Direct* 2012;**7**:43.
- Dalmaso M, Aubert J, Briard-Bion V *et al.* A Temporal -omic Study of *Propionibacterium freudenreichii* CIRM-BIA1T Adaptation Strategies in Conditions Mimicking Cheese Ripening in the Cold. Gilbert JA (ed.). *PLoS ONE* 2012a;**7**:e29083.
- Dalmaso M, Aubert J, Briard-Bion V *et al.* Correction: A Temporal -omic Study of *Propionibacterium freudenreichii* CIRM-BIA1T Adaptation Strategies in Conditions Mimicking Cheese Ripening in the Cold. *PLoS ONE* 2012b;**7**:10.1371/annotation/e0ff065d – a52d – 44f2–8727 – 328393ed60b6.
- Dalmaso M, Nicolas P, Falentin H *et al.* Multilocus sequence typing of *Propionibacterium freudenreichii*. *Int J Food Microbiol* 2011;**145**:113–20.
- Davis C. Enumeration of probiotic strains: Review of culture-dependent and alternative techniques to quantify viable bacteria. *J Microbiol Methods* 2014;**103**:9–17.
- Del Chierico F, Ancora M, Marcacci M *et al.* Choice of next-generation sequencing pipelines. *Bacterial Pangenomics*. Springer, 2015, 31–47.
- Derveaux S, Vandesompele J, Hellemans J. How to do successful gene expression analysis using real-time PCR. *Methods* 2010;**50**:227–30.
- de Sá PH, Veras AA, Carneiro AR *et al.* The impact of quality filter for RNA-Seq. *Gene* 2015;**563**:165–71.

- Dherbécourt J, Falentin H, Canaan S *et al.* A genomic search approach to identify esterases in *Propionibacterium freudenreichii* involved in the formation of flavour in Emmental cheese. *Microb Cell Factories* 2008;**7**:16.
- Dugar G, Herbig A, Förstner KU *et al.* High-resolution transcriptome maps reveal strain-specific regulatory features of multiple *Campylobacter jejuni* isolates. 2013.
- EFSA experts. The maintenance of the list of QPS microorganisms intentionally added to food or feed - Scientific Opinion of the Panel on Biological Hazards. *EFSA Journal*. 2008.
- Errington J. *Bacillus subtilis* sporulation: regulation of gene expression and control of morphogenesis. *Microbiol Rev* 1993;**57**:1–33.
- Faith JJ, Guruge JL, Charbonneau M *et al.* The long-term stability of the human gut microbiota. *Science* 2013;**341**:1237439.
- Falentin H, Deutsch S-M, Jan G *et al.* The Complete Genome of *Propionibacterium freudenreichii* CIRM-BIA1T, a Hardy Actinobacterium with Food and Probiotic Applications. Ahmed N (ed.). *PLoS ONE* 2010a;**5**:e11748.
- Falentin H, Deutsch S-M, Jan G *et al.* The Complete Genome of *Propionibacterium freudenreichii* CIRM-BIA1T, a Hardy Actinobacterium with Food and Probiotic Applications. *PLoS ONE* 2010b;**5**:e11748.
- Falentin H, Postollec F, Parayre S *et al.* Specific metabolic activity of ripening bacteria quantified by real-time reverse transcription PCR throughout Emmental cheese manufacture. *Int J Food Microbiol* 2010c;**144**:10–9.
- Fatica A, Bozzoni I. Long non-coding RNAs: new players in cell differentiation and development. *Nat Rev Genet* 2014;**15**:7–21.
- Fine AE, Bolin CA, Gardiner JC *et al.* A Study of the Persistence of *Mycobacterium bovis* in the Environment under Natural Weather Conditions in Michigan, USA, A Study of the Persistence of *Mycobacterium bovis* in the Environment under Natural Weather Conditions in Michigan, USA. *Vet Med Int Vet Med Int* 2011;**2011**, **2011**:e765430.
- Finkel SE. Long-term survival during stationary phase: evolution and the GASP phenotype. *Nat Rev Microbiol* 2006a;**4**:113–20.
- Finkel SE. Long-term survival during stationary phase: evolution and the GASP phenotype. *Nat Rev Microbiol* 2006b;**4**:113–20.
- Florent. Vitamins. *Biotechnology*. 1986:119–58.
- Förstner KU, Vogel J, Sharma CM. READemption—A tool for the computational analysis of deep-sequencing-based transcriptome data. *Bioinformatics* 2014:btu533.

- Fouquier d'Herouel A, Wessner F, Halpern D *et al.* A simple and efficient method to search for selected primary transcripts: non-coding and antisense RNAs in the human pathogen *Enterococcus faecalis*. *Nucleic Acids Res* 2011;**39**:e46–e46.
- Garber M, Grabherr MG, Guttman M *et al.* Computational methods for transcriptome annotation and quantification using RNA-seq. *Nat Methods* 2011;**8**:469–77.
- García Hernández F, Pérez Mendoza JL. [Fermentation of agro-industrial wastes for obtaining metabolites of industrial interest]. *Bol Estud Méd Biológicos* 1983;**32**:177–95.
- Gasiunas G, Barrangou R, Horvath P *et al.* Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci* 2012;**109**:E2579–86.
- Gentleman RC, Carey VJ, Bates DM *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 2004;**5**:R80.
- Geuking MB, Köller Y, Rupp S *et al.* The interplay between the gut microbiota and the immune system. *Gut Microbes* 2014;**5**:411–8.
- Gómez-Lozano M, Marvig RL, Molin S *et al.* Genome-wide identification of novel small RNAs in *Pseudomonas aeruginosa*. *Environ Microbiol* 2012;**14**:2006–16.
- Gophna U, Ron EZ. Virulence and the heat shock response. *Int J Med Microbiol IJMM* 2003;**292**:453–61.
- Gordon JI, Faith JJ. *METHODS OF LOW ERROR AMPLICON SEQUENCING (LEA-Seq) AND THE USE THEREOF*. Google Patents, 2014.
- Gottesman S, Storz G. Bacterial Small RNA Regulators: Versatile Roles and Rapidly Evolving Variations. *Cold Spring Harb Perspect Biol* 2011;**3**, DOI: 10.1101/cshperspect.a003798.
- Grainger DC, Aiba H, Hurd D *et al.* Transcription factor distribution in *Escherichia coli*: studies with FNR protein. *Nucleic Acids Res* 2007;**35**:269–78.
- Greenblatt CL, Baum J, Klein BY *et al.* *Micrococcus luteus* - Survival in Amber. *Microb Ecol* 2004;**48**:120–7.
- Haas BJ, Chin M, Nusbaum C *et al.* How deep is deep enough for RNA-Seq profiling of bacterial transcriptomes? *BMC Genomics* 2012;**13**:734.
- Haas BJ, Papanicolaou A, Yassour M *et al.* De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc* 2013;**8**:1494–512.
- Hafner M, Landthaler M, Burger L *et al.* Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 2010;**141**:129–41.

- Hague A, Elder DJ, Hicks DJ *et al.* Apoptosis in colorectal tumour cells: induction by the short chain fatty acids butyrate, propionate and acetate and by the bile salt deoxycholate. *Int J Cancer J Int Cancer* 1995;**60**:400–6.
- Hamon E, Horvatovich P, Bisch M *et al.* Investigation of Biomarkers of Bile Tolerance in *Lactobacillus casei* Using Comparative Proteomics. *J Proteome Res* 2012;**11**:109–18.
- Hannon GJ, Rossi JJ. Unlocking the potential of the human genome with RNA interference. *Nature* 2004;**431**:371–8.
- Harrington ED, Jensen LJ, Bork P. Predicting biological networks from genomic data. *FEBS Lett* 2008;**582**:1251–8.
- Havens MA, Duelli DM, Hastings ML. Targeting RNA splicing for disease therapy. *Wiley Interdiscip Rev RNA* 2013;**4**:247–66.
- Haycocks JRJ, Sharma P, Stringer AM *et al.* The Molecular Basis for Control of ETEC Enterotoxin Expression in Response to Environment and Host. *PLoS Pathog* 2015;**11**:e1004605.
- Heidrich N, Dugar G, Vogel J *et al.* Investigating CRISPR RNA Biogenesis and Function Using RNA-seq. *CRISPR Methods Protoc* 2015:1–21.
- Herbig A, Nieselt K. nocoRNAC: characterization of non-coding RNAs in prokaryotes. *BMC Bioinformatics* 2011;**12**:40.
- Hervé C, Fondrevez M, Chéron A *et al.* Transcarboxylase mRNA: a marker which evidences *P. freudenreichii* survival and metabolic activity during its transit in the human gut. *Int J Food Microbiol* 2007;**113**:303–14.
- Hilker R, Stadermann KB, Doppmeier D *et al.* ReadXplorer—visualization and analysis of mapped sequences. *Bioinformatics* 2014:btu205.
- Hooper LV, Macpherson AJ. Immune adaptations that maintain homeostasis with the intestinal microbiota. *Nat Rev Immunol* 2010;**10**:159–69.
- Huang Y, Adams MC. In vitro assessment of the upper gastrointestinal tolerance of potential probiotic dairy propionibacteria. *Int J Food Microbiol* 2004;**91**:253–60.
- Humphrys MS, Creasy T, Sun Y *et al.* Simultaneous Transcriptional Profiling of Bacteria and Their Host Cells. *PLoS ONE* 2013;**8**:e80597.
- Innocenti N, Golumbeanu M, d’Hérouël AF *et al.* Whole genome mapping of 5’RNA ends in bacteria by tagged sequencing: A comprehensive view in *Enterococcus faecalis*. *ArXiv Prepr ArXiv14101925* 2014.
- Isabella VM, Clark VL. Deep sequencing-based analysis of the anaerobic stimulon in *Neisseria gonorrhoeae*. *BMC Genomics* 2011;**12**:51.

- Isawa K, HOJO K, YODAN *et al.* Isolation and Identification of a New Bifidogenic Growth Stimulator Produced by *Propionibacterium freudenreichii* ET-3. *Biosci Biotechnol Biochem* 2002;**66**:679–81.
- Ishino Y, Shinagawa H, Makino K *et al.* Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol* 1987;**169**:5429–33.
- Jan G, Belzacq A-S, Haouzi D *et al.* Propionibacteria induce apoptosis of colorectal carcinoma cells via short-chain fatty acids acting on mitochondria. *Cell Death Differ* 2002;**9**:179–88.
- Jan G, Leverrier P, Pichereau V *et al.* Changes in Protein Synthesis and Morphology during Acid Adaptation of *Propionibacterium freudenreichii*. *Appl Environ Microbiol* 2001;**67**:2029–36.
- Jan G, Rouault A, Maubois J-L. Acid stress susceptibility and acid adaptation of *Propionibacterium freudenreichii* subsp. *shermanii*. *Le Lait* 2000;**80**:325–36.
- Jansen R, Embden J, Gaastra W *et al.* Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* 2002;**43**:1565–75.
- Jinek M, Chylinski K, Fonfara I *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 2012;**337**:816–21.
- Johnson DS, Mortazavi A, Myers RM *et al.* Genome-wide mapping of in vivo protein-DNA interactions. *Science* 2007;**316**:1497–502.
- Kahramanoglou C, Seshasayee ASN, Prieto AI *et al.* Direct and indirect effects of H-NS and Fis on global gene expression control in *Escherichia coli*. *Nucleic Acids Res* 2011;**39**:2073–91.
- Kalia VC. *Quorum Sensing vs Quorum Quenching: A Battle with No End in Sight*. Springer, 2014.
- Karimi P, Islami F, Anandasabapathy S *et al.* Gastric Cancer: Descriptive Epidemiology, Risk Factors, Screening, and Prevention. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol* 2014;**23**:700–13.
- Kato M, Chen X, Inukai S *et al.* Age-associated changes in expression of small, noncoding RNAs, including microRNAs, in *C. elegans*. *Rna* 2011;**17**:1804–20.
- Kato M, Slack FJ. Ageing and the small, non-coding RNA world. *Ageing Res Rev* 2013;**12**:429–35.
- Koboldt DC, Steinberg KM, Larson DE *et al.* The Next-Generation Sequencing Revolution and Its Impact on Genomics. *Cell* 2013;**155**:27–38.
- Kocevar N, Hudler P, Komel R. The progress of proteomic approaches in searching for cancer biomarkers. *New Biotechnol* 2013;**30**:319–26.

- Koussémon M, Combet-Blanc Y, Patel BK *et al.* Propionibacterium microaerophilum sp. nov., a microaerophilic bacterium isolated from olive mill wastewater. *Int J Syst Evol Microbiol* 2001;**51**:1373–82.
- Kratz A, Carninci P. The devil in the details of RNA-seq. *Nat Biotechnol* 2014;**32**:882–4.
- Kröger C, Dillon SC, Cameron ADS *et al.* The transcriptional landscape and small RNAs of Salmonella enterica serovar Typhimurium. *Proc Natl Acad Sci* 2012;**109**:E1277–86.
- Kulaev IS, Vagabov VM. Polyphosphate metabolism in micro-organisms. *Adv Microb Physiol* 1983;**24**:83–171.
- Kunin V, Sorek R, Hugenholtz P. Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol* 2007;**8**:R61.
- Kusano K, Yamada H, Niwa M *et al.* Propionibacterium cyclohexanicum sp. nov., a new acid-tolerant omega-cyclohexyl fatty acid-containing propionibacterium isolated from spoiled orange juice. *Int J Syst Bacteriol* 1997;**47**:825–31.
- Kusumoto A, Miyashita M, Kawamoto K. Deletion in the C-terminal domain of ClpX delayed entry of Salmonella enterica into a viable but non-culturable state. *Res Microbiol* 2013;**164**:335–41.
- Lan A, Bruneau A, Bensaada M *et al.* Increased induction of apoptosis by Propionibacterium freudenreichii TL133 in colonic mucosal crypts of human microbiota-associated rats treated with 1,2-dimethylhydrazine. *Br J Nutr* 2008;**100**:1251–9.
- Lan A, Bruneau A, Philippe C *et al.* Survival and metabolic activity of selected strains of Propionibacterium freudenreichii in the gastrointestinal tract of human microbiota-associated rats. *Br J Nutr* 2007;**97**:714.
- Langsrud and Reinbold. Flavor development and microbiology of Swiss cheese- A review. III. Ripening and flavor production. *J Milk Food Technol* 1973;**36**:593–609.
- Le Boucher C, Courant F, Jeanson S *et al.* First mass spectrometry metabolic fingerprinting of bacterial metabolism in a model cheese. *Food Chem* 2013;**141**:1032–40.
- Lee JT, Bartolomei MS. X-inactivation, imprinting, and long noncoding RNAs in health and disease. *Cell* 2013;**152**:1308–23.
- Leimena MM, Ramiro-Garcia J, Davids M *et al.* A comprehensive metatranscriptome analysis pipeline and its validation using human small intestine microbiota datasets. *BMC Genomics* 2013;**14**:530.
- Le Maréchal C, Peton V, Plé C *et al.* Surface proteins of Propionibacterium freudenreichii are involved in its anti-inflammatory properties. *J Proteomics* 2015;**113**:447–61.

- Lemunier M, Francou C, Rousseaux S *et al.* Long-term survival of pathogenic and sanitation indicator bacteria in experimental biowaste composts. *Appl Environ Microbiol* 2005;**71**:5779–86.
- Leverrier P, Dimova D, Pichereau V *et al.* Susceptibility and adaptive response to bile salts in *Propionibacterium freudenreichii*: physiological and proteomic analysis. *Appl Environ Microbiol* 2003a;**69**:3809–18.
- Leverrier P, Dimova D, Pichereau V *et al.* Susceptibility and Adaptive Response to Bile Salts in *Propionibacterium freudenreichii*: Physiological and Proteomic Analysis. *Appl Environ Microbiol* 2003b;**69**:3809–18.
- Lewis K. Programmed Death in Bacteria. *Microbiol Mol Biol Rev* 2000;**64**:503–14.
- Lilly DM, Stillwell RH. PROBIOTICS: GROWTH-PROMOTING FACTORS PRODUCED BY MICROORGANISMS. *Science* 1965;**147**:747–8.
- Li S, Tighe SW, Nicolet CM *et al.* Multi-platform assessment of transcriptome profiling using RNA-seq in the ABRF next-generation sequencing study. *Nat Biotechnol* 2014;**32**:915–25.
- Liu Y, Zhou J, White KP. RNA-seq differential expression studies: more sequence or more replication? *Bioinformatics* 2014;**30**:301–4.
- Loux L, Mariadassou M, Almeida S *et al.* Mutations and genomic islands explain the strain dependency of sugar utilization among 21 strains of *Propionibacterium freudenreichii*. *BMC Genomics* 2015a.
- Loux V, Mariadassou M, Almeida S *et al.* Mutations and genomic islands can explain the strain dependency of sugar utilization in 21 strains of *Propionibacterium freudenreichii*. *BMC Genomics* 2015b;**16**:296.
- Luerce T, Gomes-Santos A, Rocha C *et al.* Anti-inflammatory effects of *Lactococcus lactis* NCDO 2118 during the remission period of chemically induced colitis. *Gut Pathog* 2014;**6**:33.
- Luijk NV, Stierli MP, Schwenninger SM *et al.* Genetics and molecular biology of propionibacteria. *Le Lait* 2002;**82**:14.
- Macklaim JM, Fernandes AD, Bella JMD *et al.* Comparative meta-RNA-seq of the vaginal microbiota and differential expression by *Lactobacillus iners* in health and dysbiosis. *Microbiome* 2013;**1**:12.
- Malik AC, Reinbold GW, Vedamuthu ER. Evaluation of the taxonomy of the *Propionibacterium*. *Canadian Journal of Microbiology*. 1968:1185–91.
- Mamanova L, Andrews RM, James KD *et al.* FRT-seq: amplification-free, strand-specific transcriptome sequencing. *Nat Methods* 2010;**7**:130–2.
- Mamanova L, Turner DJ. Low-bias, strand-specific transcriptome Illumina sequencing by on-flowcell reverse transcription (FRT-seq). *Nat Protoc* 2011;**6**:1736–47.



- Mantere-Alhonen S. Die Propionibakterien der Molkereiindustrie als Darmkanalmikroben. *Meijeritiet Aikak* 1982;**40**.
- Mantere-Alhonen S. Propionibacteria used as probiotics - A review. *Le Lait* 1995;**75**:447–52.
- Mantere-Alhonen S, Mâkinen E. A new sour milk product containing *Propionibacterium freudenreichii* and *Lactobacillus acidophilus*. *Meijeritiet Aikak* 1986;**40**:49–58.
- Margulies M, Egholm M, Altman WE *et al*. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 2005;**437**:376–80.
- Martin JA, Wang Z. Next-generation transcriptome assembly. *Nat Rev Genet* 2011;**12**:671–82.
- Marwaha SS, Sethi RP, Kennedy JF. Influence of 5,6-dimethylbenzimidazole (DMB) on vitamin B12 biosynthesis by strains of *Propionibacterium*. *Enzyme Microb Technol* 1983;**5**:361–4.
- McClure R, Balasubramanian D, Sun Y *et al*. Computational analysis of bacterial RNA-Seq data. *Nucleic Acids Res* 2013a:gkt444.
- McClure R, Balasubramanian D, Sun Y *et al*. Computational analysis of bacterial RNA-Seq data. *Nucleic Acids Res* 2013b;**41**:e140–e140.
- McGettigan PA. Transcriptomics in the RNA-seq era. *Curr Opin Chem Biol* 2013;**17**:4–11.
- Meile L, Dasen G, Miescher S *et al*. Classification of propionic acid bacteria and approaches to applied genetics. *Le Lait* 1999;**79**:71–8.
- Meurice G, Jacob D, Deborde C *et al*. Whole genome sequencing project of a dairy *Propionibacterium freudenreichii* subsp. *shermanii* genome: progress and first bioinformatic analysis. *Le Lait* 2004;**84**:15–24.
- Meyrand M, Guillot A, Goin M *et al*. Surface Proteome Analysis of a Natural Isolate of *Lactococcus lactis* Reveals the Presence of Pili Able to Bind Human Intestinal Epithelial Cells. *Mol Cell Proteomics* 2013;**12**:3935–47.
- Mogensen G, SALMINEN;S, O'BRIEN;J *et al*. *Inventory of Microorganisms with a Documented History of Use in Food*. Brussels, BELGIQUE: International Dairy Federation, 2002.
- Moore W, Cato E. VALIDITY OF PROPIONIBACTERIUM ACNES (GILCHRIST) DOUGLAS AND GUNTER COMB. NOV. *Journal of Bacteriology*. 1963:870–4.
- Mulder NJ, Akinola RO, Mazandu GK *et al*. Using biological networks to improve our understanding of infectious diseases. *Comput Struct Biotechnol J* 2014;**11**:1–10.

- Murooka Y, Piao Y, Kiatpapan P *et al.* Production of tetrapyrrole compounds and vitamin B12 using genetically engineering of *Propionibacterium freudenreichii*. An overview. *Le Lait* 2005;**85**:14.
- Myers KS, Park DM, Beauchene NA *et al.* Defining bacterial regulons using ChIP-seq methods. *Methods* 2015.
- Myers KS, Yan H, Ong IM *et al.* Genome-scale analysis of *Escherichia coli* FNR reveals complex features of transcription factor binding. *PLoS Genet* 2013;**9**:e1003565.
- Navarro E, Serrano-Heras G, Castaño MJ *et al.* Real-time PCR detection chemistry. *Clin Chim Acta* 2015;**439**:231–50.
- Ndimba BK, Ndimba RJ, Johnson TS *et al.* Biofuels as a sustainable energy source: an update of the applications of proteomics in bioenergy crops and algae. *J Proteomics* 2013;**93**:234–44.
- Nelson DL, Nelson DL, Lehninger AL *et al.* *Lehninger Principles of Biochemistry*. New York: W.H. Freeman, 2008.
- Ngounou Wetie AG, Sokolowska I, Woods AG *et al.* Protein–protein interactions: switch from classical methods to proteomics and bioinformatics-based approaches. *Cell Mol Life Sci* 2014;**71**:205–28.
- Okada Y, Tsuzuki Y, Miyazaki J *et al.* *Propionibacterium freudenreichii* component 1,4-dihydroxy-2-naphthoic acid (DHNA) attenuates dextran sodium sulphate induced colitis by modulation of bacterial flora and lymphocyte homing. *Gut* 2006;**55**:681–8.
- Okada Y, Tsuzuki Y, Narimatsu K *et al.* 1,4-Dihydroxy-2-naphthoic acid from *Propionibacterium freudenreichii* reduces inflammation in interleukin-10-deficient mice with colitis by suppressing macrophage-derived proinflammatory cytokines. *J Leukoc Biol* 2013;**94**:473–80.
- Olszewska M, Staniewski B, Łaniewska-Trokenheim Ł. Cell Viability of *Bifidobacterium lactis* Strain in Long-Term Storage Butter Assessed with the Plate Count and Fluorescence Techniques. *Czech J Food Sci* 2012;**30**:421–8.
- Orla-Jensen. *Die Hauptlinien des natürlichen Bakteriensystems nebst einer Uebersicht der Gärungsphänomene.*, 1909.
- Osmundson J, Dewell S, Darst SA. RNA-Seq Reveals Differential Gene Expression in *Staphylococcus aureus* with Single-Nucleotide Resolution. *PLoS ONE* 2013;**8**:e76572.
- Otto A, Becher D, Schmidt F. Quantitative proteomics in the field of microbiology. *PROTEOMICS* 2014;**14**:547–65.
- Ouwehand AC, Kirjavainen PV, Grönlund M-M *et al.* Adhesion of probiotic microorganisms to intestinal mucus. *Int Dairy J* 1999;**9**:623–30.

- Ouwehand AC, Suomalainen T, Tölkö S *et al.* In vitro adhesion of propionic acid bacteria to human intestinal mucus. *Le Lait* 2002;**82**:8.
- Ouwehand AC, Tiihonen K, Saarinen M *et al.* Influence of a combination of *Lactobacillus acidophilus* NCFM and lactitol on healthy elderly: intestinal and immune parameters. *Br J Nutr* 2009;**101**:367–75.
- Papenfors K, Vogel J. Regulatory RNA in Bacterial Pathogens. *Cell Host Microbe* 2010;**8**:116–27.
- Parayre S, Falentin H, Madec M-N *et al.* Easy DNA extraction method and optimisation of PCR-Temporal Temperature Gel Electrophoresis to identify the predominant high and low GC-content bacteria from dairy products. *J Microbiol Methods* 2007;**69**:431–41.
- Parizzi LP, Grassi MCB, Llerena LA *et al.* The genome sequence of *Propionibacterium acidipropionici* provides insights into its biotechnological and industrial potential. *BMC Genomics* 2012;**13**:562.
- Park PJ. ChIP–seq: advantages and challenges of a maturing technology. *Nat Rev Genet* 2009;**10**:669–80.
- Parvez S, Malik KA, Ah Kang S *et al.* Probiotics and their fermented food products are beneficial for health. *J Appl Microbiol* 2006;**100**:1171–85.
- Passalacqua KD, Varadarajan A, Ondov BD *et al.* Structure and Complexity of a Bacterial Transcriptome. *J Bacteriol* 2009;**191**:3203–11.
- Patenge N, Pappesch R, Khani A *et al.* Genome-wide analyses of small non-coding RNAs in streptococci. *Front Genet* 2015;**6**:189.
- Pérez Chaia A, Zarate G. Dairy propionibacteria from milk or cheese diets remain viable and enhance propionic acid production in the mouse cecum. *Le Lait* 2005;**85**:85–98.
- Pérez Chaia A, Zarate G, Oliver G. The probiotic properties of propionibacteria. *Le Lait* 1999;**79**:175–85.
- Perkins TT, Kingsley RA, Fookes MC *et al.* A strand-specific RNA-Seq analysis of the transcriptome of the typhoid bacillus *Salmonella typhi*. *PLoS Genet* 2009;**5**:e1000569.
- Piao Y, Yamashita M, Kawaraichi N *et al.* Production of vitamin B12 in genetically engineered *Propionibacterium freudenreichii*. *J Biosci Bioeng* 2004;**98**:167–73.
- Picard C, Fioramonti J, Francois A *et al.* Review article: bifidobacteria as probiotic agents – physiological effects and clinical benefits. *Aliment Pharmacol Ther* 2005;**22**:495–512.
- Pinto AC, Melo-Barbosa HP, Miyoshi A *et al.* Review Application of RNA-seq to reveal the transcript profile in bacteria. *Genet Mol Res* 2011;**10**:1707–18.

- Pinto AC, Ramos RTJ, Silva WM *et al.* The core stimulon of *Corynebacterium pseudotuberculosis* strain 1002 identified using ab initio methodologies. *Integr Biol* 2012;**4**:789.
- Pinto AC, Sá PHCG de, Ramos RTJ *et al.* Differential transcriptional profile of *Corynebacterium pseudotuberculosis* in response to abiotic stresses. *BMC Genomics* 2014;**15**:14.
- Poonam, Pophaly SD, Tomar SK *et al.* Multifaceted attributes of dairy propionibacteria: a review. *World J Microbiol Biotechnol* 2012;**28**:3081–95.
- Primm TP, Andersen SJ, Mizrahi V *et al.* The Stringent Response of *Mycobacterium tuberculosis* Is Required for Long-Term Survival. *J Bacteriol* 2000;**182**:4889–98.
- Quigley ME, Hudson GJ, Englyst HN. Determination of resistant short-chain carbohydrates (non-digestible oligosaccharides) using gas–liquid chromatography. *Food Chem* 1999;**65**:381–90.
- Ramos RT, Carneiro AR, Baumbach J *et al.* Analysis of quality raw data of second generation sequencers with Quality Assessment Software. *BMC Res Notes* 2011;**4**:130.
- Rao VS, Srinivas K, Sujini GN *et al.* Protein-Protein Interaction Detection: Methods and Analysis. *Int J Proteomics* 2014;**2014**:1–12.
- Rittershaus ESC, Baek S, Sasseti CM. The Normalcy of Dormancy. *Cell Host Microbe* 2013;**13**:643–51.
- Robertson G, Hirst M, Bainbridge M *et al.* Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat Methods* 2007;**4**:651–7.
- Sabin LR, Delás MJ, Hannon GJ. Dogma Derailed: The Many Influences of RNA on the Genome. *Mol Cell* 2013a;**49**:783–94.
- Sabin LR, Delás MJ, Hannon GJ. Dogma Derailed: The Many Influences of RNA on the Genome. *Mol Cell* 2013b;**49**:783–94.
- Sanchez B, Champomier-Verges M-C, Stuer-Lauridsen B *et al.* Adaptation and Response of *Bifidobacterium animalis* subsp. *lactis* to Bile: a Proteomic and Physiological Approach. *Appl Environ Microbiol* 2007;**73**:6757–67.
- Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 1977;**74**:5463–7.
- Santos Rocha C, Lakhdari O, Blottière HM *et al.* Anti-inflammatory properties of dairy lactobacilli: *Inflamm Bowel Dis* 2012;**18**:657–66.
- Saraoui T, Parayre S, Guerne G *et al.* A unique in vivo experimental approach reveals metabolic adaptation of the probiotic *Propionibacterium freudenreichii* to the colon environment. *BMC Genomics* 2013a;**14**:911.

- Saraoui T, Parayre S, Guernec G *et al.* A unique in vivo experimental approach reveals metabolic adaptation of the probiotic *Propionibacterium freudenreichii* to the colon environment. *BMC Genomics* 2013b;**14**:911.
- Sartor RB. Microbial Influences in Inflammatory Bowel Diseases. *Gastroenterology* 2008;**134**:577–94.
- Sathish JG, Sethu S, Bielsky M-C *et al.* Challenges and approaches for the development of safer immunomodulatory biologics. *Nat Rev Drug Discov* 2013;**12**:306–24.
- Sendler E, Johnson GD, Krawetz SA. Local and global factors affecting RNA sequencing analysis. *Anal Biochem* 2011;**419**:317–22.
- Syednasrollah F, Laiho A, Elo LL. Comparison of software packages for detecting differential expression in RNA-seq studies. *Brief Bioinform* 2013, DOI: 10.1093/bib/bbt086.
- Sharma CM, Hoffmann S, Darfeuille F *et al.* The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* 2010;**464**:250–5.
- Singh SS, Singh N, Bonocora RP *et al.* Widespread suppression of intragenic transcription initiation by H-NS. *Genes Dev* 2014;**28**:214–9.
- Sittka A, Lucchini S, Papenfort K *et al.* Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLoS Genet* 2008;**4**:e1000163.
- Sorek R, Cossart P. Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nat Rev Genet* 2010;**11**:9–16.
- Storz G, Vogel J, Wassarman KM. Regulation by Small RNAs in Bacteria: Expanding Frontiers. *Mol Cell* 2011;**43**:880–91.
- Stowers CC, Cox BM, Rodriguez BA. Development of an industrializable fermentation process for propionic acid production. *J Ind Microbiol Biotechnol* 2014;**41**:837–52.
- Stringer AM, Currenti S, Bonocora RP *et al.* Genome-scale analyses of *Escherichia coli* and *Salmonella enterica* AraC reveal noncanonical targets and an expanded core regulon. *J Bacteriol* 2014;**196**:660–71.
- Suomalainen T, Lagström H, Mättö J *et al.* Influence of whey-based fruit juice containing *Lactobacillus rhamnosus* on intestinal well-being and humoral immune response in healthy adults. *LWT - Food Sci Technol* 2006;**39**:788–95.
- Thierry A, Deutsch S-M, Falentin H *et al.* New insights into physiology and metabolism of *Propionibacterium freudenreichii*. *Int J Food Microbiol* 2011;**149**:19–27.
- Thierry A, Maillard M-B, Richoux R *et al.* *Propionibacterium freudenreichii* strains quantitatively affect production of volatile compounds in Swiss cheese. *Le Lait* 2005;**85**:18.

- Tjaden B. De novo assembly of bacterial transcriptomes from RNA-seq data. *Genome Biol* 2015;**16**:1.
- Vallenet D, Nordmann P, Barbe V *et al.* Comparative Analysis of Acinetobacters: Three Genomes for Three Lifestyles. Hall N (ed.). *PLoS ONE* 2008;**3**:e1805.
- Vandamme E. Biotechnology of vitamins, pigments and growth factors. *Elsevier*. 1989:257–84.
- Vandamme EJ. Production of vitamins, coenzymes and related biochemicals by biotechnological processes. *J Chem Technol Biotechnol* 1992;**53**:313–27.
- van Dijk EL, Auger H, Jaszczyszyn Y *et al.* Ten years of next-generation sequencing technology. *Trends Genet* 2014a;**30**:418–26.
- van Dijk EL, Auger H, Jaszczyszyn Y *et al.* Ten years of next-generation sequencing technology. *Trends Genet* 2014b;**30**:418–26.
- Vangay P, Ward T, Gerber JS *et al.* Antibiotics, Pediatric Dysbiosis, and Disease. *Cell Host Microbe* 2015;**17**:553–64.
- Van Verk MC, Hickman R, Pieterse CM *et al.* RNA-Seq: revelation of the messengers. *Trends Plant Sci* 2013;**18**:175–9.
- Van Wyk J, Witthuhn RC, Britz TJ. Optimisation of vitamin B12 and folate production by *Propionibacterium freudenreichii* strains in kefir. *Int Dairy J* 2011;**21**:69–74.
- Velasquez-Manoff M. Gut Microbiome: The Peacekeepers. *Nature* 2015;**518**:S3–11.
- Vergara-Irigaray M, Fookes MC, Thomson NR *et al.* RNA-seq analysis of the influence of anaerobiosis and FNR on *Shigella flexneri*. *BMC Genomics* 2014;**15**:438.
- von Freudenreich and Orla-Jensen. Über die in emmentalerkäse stattfindene propionsäure-gärung. *Bakteriol* 1906;**17**:529–46.
- Vorobjeva LI. *Propionibacteria*. *Kluwer Acad Publ* 1999.
- de Vos WM, de Vos EAJ. Role of the intestinal microbiome in health and disease: from correlation to causation. *Nutr Rev* 2012;**70 Suppl 1**:S45–56.
- Wade JT, Struhl K, Busby SJ *et al.* Genomic analysis of protein–DNA interactions in bacteria: insights into transcription and chromosome organization. *Mol Microbiol* 2007;**65**:21–6.
- Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. *Mol Cell* 2011;**43**:904–14.
- Wang P, Zhang Z, Jiao Y *et al.* Improved propionic acid and 5,6-dimethylbenzimidazole control strategy for vitamin B12 fermentation by *Propionibacterium freudenreichii*. *J Biotechnol* 2015;**193**:123–9.

- Wang Z, Yang S-T. Propionic acid production in glycerol/glucose co-fermentation by *Propionibacterium freudenreichii* subsp. *shermanii*. *Bioresour Technol* 2013a;**137**:116–23.
- Wang Z, Yang S-T. Propionic acid production in glycerol/glucose co-fermentation by *Propionibacterium freudenreichii* subsp. *shermanii*. *Bioresour Technol* 2013b;**137**:116–23.
- Warminska-Radyko I, Laniewska-Moroz L, Babuchowski A. Possibilities for stimulation of *Bifidobacterium* growth by propionibacteria. *Le Lait* 2002;**82**:10.
- Warren AS, Aurrecoechea C, Brunk B *et al.* RNA-Rocket: an RNA-Seq analysis resource for infectious disease research. *Bioinformatics* 2015:btv002.
- Wen J, Deng X, Li Z *et al.* Transcriptomic Response of *Listeria monocytogenes* during the Transition to the Long-Term-Survival Phase. *Appl Environ Microbiol* 2011;**77**:5966–72.
- Westermann AJ. 1, Gorski, SA and Vogel, J.(2012) Dual RNA-seq of pathogen and host. *Nat Rev Microbiol* **10**:618–30.
- Wheat WH, Casali AL, Thomas V *et al.* Long-term survival and virulence of *Mycobacterium leprae* in amoebal cysts. *PLoS Negl Trop Dis* 2014;**8**:e3405.
- Wiegand S, Dietrich S, Hertel R *et al.* RNA-Seq of *Bacillus licheniformis*: active regulatory RNA features expressed within a productive fermentation. *BMC Genomics* 2013;**14**:667.
- Wood HG. Metabolic cycles in the fermentation by propionic acid bacteria. *Curr Top Cell Regul* 1981;**18**:255–87.
- Yan Y, Su S, Meng X *et al.* Determination of sRNA Expressions by RNA-seq in *Yersinia pestis* Grown In Vitro and during Infection. *PLoS ONE* 2013;**8**:e74495.
- Ye K, Shijo M, Miyano K *et al.* Metabolic Pathway of *Propionibacterium* Growing with Oxygen: Enzymes, <sup>13</sup>C NMR Analysis, and Its Application for Vitamin B12 Production with Periodic Fermentation. *Biotechnol Prog* 1999;**15**:201–7.
- Zarate G, Chaia AP, Oliver G. Some characteristics of practical relevance of the beta-galactosidase from potential probiotic strains of *Propionibacterium acidipropionici*. *Anaerobe* 2002;**8**:259–67.
- Zárate G, Morata De Ambrosini V, Perez Chaia A *et al.* Some factors affecting the adherence of probiotic *Propionibacterium acidipropionici* CRL 1198 to intestinal epithelial cells. *Can J Microbiol* 2002;**48**:449–57.
- Zoepfel J, Randau L. RNA-Seq analyses reveal CRISPR RNA processing and regulation patterns. *Biochem Soc Trans* 2013;**41**:1459–63.

## **Anexos / Annexes**



**Anexo 1: Capítulo de livro: *Genomics* publicado em *A Text Book of Biotechnology* (disponível em [www.smgebooks.com](http://www.smgebooks.com))**

**Annexe 1 : Chapitre du livre : *Genomics* publié dans *A Text Book of Biotechnology* (disponible en [www.smgebooks.com](http://www.smgebooks.com))**

## **Genomics**

Flavia Figueira Aburjaile<sup>1</sup>, Mariana Passos Santana<sup>1</sup>, Marcus Vinicius Canário Viana<sup>1</sup>, Wanderson Marques Silva<sup>1</sup>, Edson Luiz Folador<sup>1,2</sup>, Artur Silva<sup>2</sup> and Vasco Azevedo<sup>1</sup>

**Aburjaile FF<sup>1</sup>, Santana MP<sup>1</sup>, Viana MVC<sup>1</sup>, Silva WM<sup>1</sup>, Folador EL<sup>1</sup>, Silva A<sup>2</sup> and Azevedo V<sup>1</sup>**

<sup>1</sup>Universidade Federal de Minas Gerais, UFMG, Belo Horizonte, Brazil

<sup>2</sup>Universidade Federal do Pará, BÉlem, Brazil

**\*Corresponding author:** Vasco Azevedo, Universidade Federal de Minas Gerais/Instituto de Ciências Biológicas, Minas Gerais, Brazil, Email: [vasco@icb.ufmg.br](mailto:vasco@icb.ufmg.br)

**Published Date:** February 15, 2015

### **Abstract**

In the last few years, the technologies have revolutionized new areas in science especially those involving genomics, proteomics and transcriptomics. This chapter approach each of these areas showing their concepts, importance and applications.

**Keywords:** genomics, structural genomics, functional genomics, transcriptomics, proteomics.

## **Genomics**

### **Structural genomics**

#### **Introduction**

Structural genomics is the analysis of sequence and structure of the genome elements such as genes, regulators and mobile elements. Sequencing is necessary to obtain this information. The identification of variants among genomes has applications in evolution studies, health, biotechnology and the comprehension of relation between genotype and phenotype. Knowledge of genome sequence is fundamental to obtain this kind of information, and sequencing is the process of characterization of nucleotide

sequence at a region (Targeted Genome Sequencing, TGS) or in the entire genome (Whole Genome Sequencing, WGS).

The first complete genome sequence of an organism was achieved in 1995, from the Gram-negative *Haemophilus influenzae*. Since then, the number of sequenced genomes is increasing, mainly because of new technologies that reduced the required cost and time, and new bioinformatics tools able to process the volume of data generated. Initiated formally in 1990, the Human Genome Project (HGP) was an international collaborative project, that sequenced the human genome in 13 years at a cost of approximately US \$ 2.7 billion (<http://www.genome.gov/12011238>).

### **Genome sequencing**

The procedure for sequencing complete genomes involves fragmentation of the DNA molecule, creation of genomic libraries, reading nucleotide sequence of each fragment and reassembles the genome by alignment of the reads. Since the 1970s, different methods have been developed to determine the nucleotide sequence of a DNA molecule.

### **Sequencing generations and technologies**

In 1977, Sanger and Nicklen published a DNA sequencing method by chain termination. This method required the creation of libraries by insertion of DNA fragments to be sequenced into cloning vectors (plasmids or artificial chromosomes) and then amplified *in vivo*. Four sequencing reactions are performed for each fragment, each one containing a proportion of one nucleotide without 3' OH group (dideoxy nucleotides), which causes termination of synthesis of a new DNA strands in random positions.

The generated fragments are separated by size on gel electrophoresis and the nucleotide sequence is determined by the last nucleotide of each fragment (Sanger, Nicklen and Coulson, 1977). Over the years, improvements have been incorporated into the method, such as *in vitro* amplification and automation by thermal cyclers, labeling the amplified fragments with fluorescent dideoxynucleotides (ddNTP's), as well as automatic sequencing based on reading fragments by capillary electrophoresis. The sequenced fragments generated by this method have a maximum size between 800 and 1000 bases. According to the National Human Genome Research Institute (NHGRI), in 2001, the cost to sequence 1 Mb (1,000,000 bases) and the complete human genome

were respectively U\$ 5,292 and U\$ 95,293,072. In 2007, the cost was reduced to U\$ 397 and U\$ 7,147 (<http://www.genome.gov/sequencingcosts/>). The amount of data generated by chain termination technology was inadequate for sequencing projects and to obtain complete genomes, even with the reduction of required costs and time.

In 2005, Next Generation Sequencing (NGS) platforms were launched, belonging to the second generation. These platforms are characterized by (i) the use of genomic libraries independent of cloning and (ii) a higher throughput compared to the first generation, by means of parallel sequencing of thousands or millions of fragments. However, the necessity of amplification in the preparation of the libraries can bias the quantification of fragments, insert sequence errors during the replication and increase the complexity of sample preparation. In addition, the smaller size of the reads and the large amount of generated data make the assembly of a more complex genome and creates a demand for computational structures with greater processing and storage capacity. In 2014, the utilization of these platforms, made the cost for sequencing 1Mb and a human genome decreased to, respectively, U\$ 102 and U\$ 3,063 million in 2007 to \$ 0.05 to U\$ 4905 (<http://www.genome.gov/sequencingcosts/>). The first released NGS technology was pyrosequencing technology. The sequencing reaction occurs in emulsion oil droplets distributed in wells of a fiber optic slide. Each oil droplet contain one bead linked to a single library fragment and PCR reagents. The reaction receives a flow of each nucleotide at a time followed by a wash step. The incorporation of a nucleotide is detected by the releasing of pyrophosphate. This technology is used in 454 Genome Sequencer platform (Roche), launched in 2005, being able to generate 200.00 reads of 110 base pairs (bp), and generating approximately 20MB of data. In 2007, Illumina/Solexa launched 1G Genetic Analyzer platform. In this technology, the fragments connected to adapters are denatured and its ends hybridized to complementary sequence adapters present on the surface of a solid blade. Complementary strand are synthesized by extension of the adapter attached to a solid plate. Also in 2007, Applied Biosystems launched SOLiD platform, based on ligation sequencing. A new complementary DNA strand is synthesized by ligation of an oligonucleotide to a primer and cleavage of oligonucleotide 3' end, releasing a fluorescent color, which represents the first two nucleotides of the incorporated oligonucleotide. The new DNA strand is removed by cleavage of the last nucleotide primer and a new cycle begins, initiated one position behind in relation to previous cycle. A color code and the sequence of each cycle colors are used to determine the

nucleotide sequence. In 2010, Ion Torrent (Applied Biosystems) released the Personal Genome Machine, which sequencing chemistry is similar to pyrosequencing, however the incorporation of nucleotides is detected by pH variation. This technology uses a semiconductor in place of fluorescence and a scanning camera, resulting in higher speed, lower cost and smaller equipment. Since then, updates of each platform were launched, in order to increase the size of fragment sequences and reduce the required time, cost and labor.

The third generation is characterized by (i) non-interruption of the sequencing reaction to detect each nucleotide incorporated (flow type of each nucleotide), (ii) larger reads and (iii) sequencing of a single molecule. This has reduced the sequencing time and simplified genome assembly in relation to the second generation, as reads sizes exceeds 1000pb. This generation technologies are based on imaging of singles molecules of DNA polymerase as they synthesize a single molecule of DNA; threading DNA polymerase to or next a nanopore and detection of nucleotide passage; or imaging of individual DNA strands by advanced microscopy. One example is the Platform II PacBio RS (Pacific Biosciences), launched in 2013, that generate reads longer than 20kb.

### **Genomic libraries**

The DNA library preparation is one of the fundamental steps for the success of a genome project. Despite the variation in protocols, the similarities are fragmentation of the sample to sizes between 50 and 500 nucleotides, selection of fragments by size, ligation to adapters. The adapters often contain elements for ligation on a solid surface, primer annealing sites for amplification and sequencing, and a barcode sequence that makes possible to sequence samples of different sources.

In single-end or fragments libraries, adapters are ligated to one fragment, generating one read. In paired-end libraries, the ends of a DNA fragment are read, generating two reads, separated by a gap of known size. In mate-pair libraries, DNA fragments are connected to an adapter, circularized, and their ends are read, generating two reads, separated by a gap of known size. The difference between paired-end and mate-pair library is circularization and larger fragments of the second one. The latter two libraries provide an approximate distance between two reads, and this information facilitates the genome assembly and makes it possible to detect genome deletions, inversions, duplications and nucleotide insertions.

## **Applications**

In studies with the genomic DNA, NGS technologies have been used in sequencing of complete genomes or fragments, analysis of genetic variation, chromatin immunoprecipitation followed by sequencing (ChIP-Seq), epigenetic markers, identification of genes associated with disease, genetic screening, monitoring of infectious agents, forensic research, metagenomics and agrigenomics. Thousands of sequencing projects and millions of genomes were released (<http://www.genomics.cn/en/index>). Among the RNA sequencing applications are analysis of gene expression and single cell transcriptome. In gene expression studies, NGS made the prior knowledge of genes unnecessary to the identification and quantification of transcripts, alternative splicing and sequence variations. The various applications of NGS led to the launch of benchtop sequencers, cheaper, faster, easier to use and accessible to small laboratories. These include GS Junior (Roche), Mi Seq (Illumina) and Personal Genome Machine (Life Technologies).

## **Genome Assembly**

After sequencing the genome is reconstructed from fragments reads, in a process called genome assembly. The assembly is performed by means of alignment and overlap of reads to generate contigs (Figure 1), and contigs overlap. It is important that all nucleotides of a genome are represented among the sequence reads. The theoretical (or expected) coverage of a genome refers to the average expected times a nucleotide is sequenced. The calculation is performed by multiplying the size of the reads by the number of reads, and dividing the result by the size of the sequenced genome. The depth of coverage refers to the number of times a particular nucleotide of the genome is represented in the reads. This number varies because the fragment library preparation, sequencing and the process of assembling the genome generate deviations in the distribution of reads. The breadth of coverage is the percentage of a given genome sequenced by a number of reads.

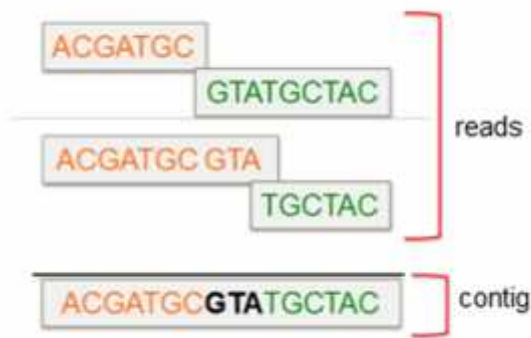


Figure 1 – Assembly of contigs sequences by reads alignment.

### Data analysis

The first assembly step is data analysis. Raw data from sequencer is converted to file formats used by assembly software. The main formats are FASTQ and FASTA. FASTA (Fast Alignment) format stores nucleotide sequences (Figure 2), while FASTQ stores nucleotide sequence and quality scores. Removal of poor quality sequences, and barcode reads is required, usually performed automatically by the assembly software. Reads quality is measured by the Phred index, a measure of error probability in identification of each nucleotide in a read.

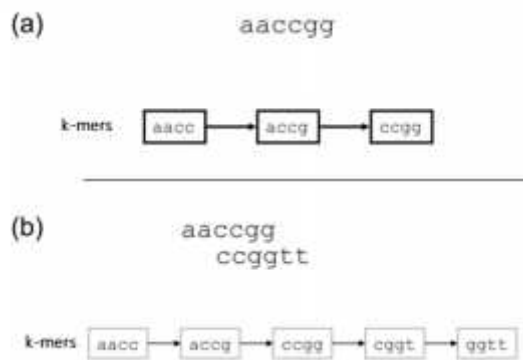
```
>Sequencia_1
ATCGATCGATGCTAGCATCGATCGATCGATCAGCTAGCTAGCTACTACGATCGATCA
GTCAGCTAGCATGCATCGATCGATCGATCGATCGATCAGCTAGCTAGCTAGCATCGA
GCTAGCTAGCATCGATCGATCGATCGATCAGTCAGCATGCATGCATCGATGCACACACACA
CACACCACACACACGTGTGTCAGCTAGGCTCGCGCGCGGCCCGTACGATCGGCCAC
ATCGATCGATGCTAGCATCGATCGATCGATCAGCTAGCTAGCTACTACGATCGATCA
>Sequencia_2
ATCGATCGATGCTAGCATCGATCGATCGATCAGCTAGCTAGCTACTACGATCGATCA
GTCAGCTAGCATGCATCGATCGATCGATCGATCGATCAGCTAGCTAGCTAGCATCGA
GCTAGCTAGCATCGATCGATCGATCGATCAGTCAGCATGCATGCATCGATGCACACACACA
ATCGATCGATGCTAGCATCGATCGATCGATCAGCTAGCTAGCTACTACGATCGATCA
GTCAGCTAGCATGCATCGATCGATCGATCGATCGATCAGCTAGCTAGCTAGCATCGA
```

Figure 2 – Fasta file containing two sequences (multifasta file).

### Assembly

To assemble a genome is necessary to find overlaps between reads and generate a consensus, or contiguous, sequence. The assembly software uses algorithms that create substrings of each read, called k-mers, and test alignments with k-mers of other reads (Figure 3) to generate contiguous sequences. The use of paired libraries facilitates the assembly process because the known distance among the pairs of readings. For example, if each paired reading belongs to a different contig, it is possible to determine

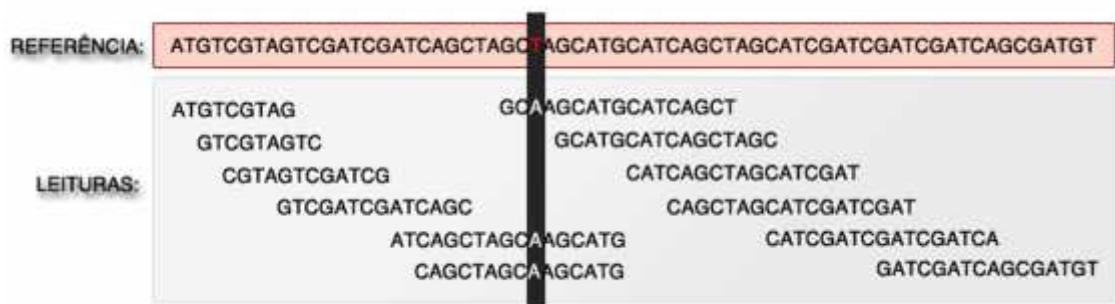
the distance and orientation between these two contigs. A genome can be assembled using de novo or ab initio methods and by reference method.



**Figure 3 – (A) k-mers of 4 nucleotides based on the read sequence “aacgg”. (B) Alignment of the reads “aacgg” and “ccggt” based on k-mers.**

The de novo assembly is performed only by reads overlapping, not requiring the use of a reference genome mapping (Figure 1). The most commonly used algorithms are the Greedy, OLC (Overlap-consensus) and De Bruijn graph. This method allows identification of absent regions in a reference genome. The difficulty of this method is repetitive regions where reads can be aligned in wrong places.

In the assembly by reference, the reads are mapped in a phylogenetically close genome. This method has lower computational cost and is useful in identifying SNPs, insertions and deletions between the reference and the new genomes (Figure 4). Another advantage is the resolution of repetitive regions, because the reads are distributed at the reference regions. However, absent sequences in the reference genome will not be mapped.



**Figure 4 – Mapping reads on a reference genome.**

### Scaffolding and closure of gaps

The contigs have to be ordered and oriented by overlaps between their ends (de novo assembly), mapping on reference genome or optical mapping. The latter is done by mapping restriction sites *in vitro* and comparison to the contigs. A contig is called scaffold after its position and orientation is defined. The not identified overlaps between the contigs and, mainly, uncovered regions cause gaps in the assembly. Gaps can be closed *in silico* by overlapping scaffold ends, extension of scaffold by paired reads libraries, mapping reads in another reference genome in the region equivalent to the gap. The closing is accomplished by *in vitro* primer design which is long to the edge of two adjacent scaffolds and subsequent amplification of the region of the gap, usually by the Sanger method by generating sequences which can cover or extend the gap ends of the scaffolds. The third-generation sequencing platforms which generate long readings are becoming obsolete closing gaps *in vitro*, mainly due to the cost required. If a genome is still contains gaps after assembly, it is called a draft. The most commonly used parameters to measure the quality of an assembly is the number of contigs, minimum and maximum contig size and N50. In a good assembly, the number of contigs and its sizes is, respectively, the minimum and the maximum possible. The N50 is the size of the contig which the sum with the others, in descending order of size, is greater than or equal to 50% of the assembled genome.

### **Genome Annotation (automatic and manual steps)**

After assembly, the structure and function of genome elements as genes, regulators and other non-coding sequences have to be identified to posterior analysis. This process is called genome annotation. Gene prediction can be made empirically, by similarity with known sequences in databases and proteins, or *ab initio*, by computational techniques, using algorithms to find stop codons, coding sequences (CDS), start codons. Normally, the algorithms to predict coding sequences are based on Markov chain models. In Prokaryotes, gene prediction is based on Open Reading Frames (ORFs) in the six possible frames. An ORF is a sequence that initiates at a start codon and finish at the nearest 3' stop codon. The prediction of false positives, incorrect start codons and overlapping genes are the main problems of this method. This problem can be minimized by searching for ribosome binding sites (RBS) to indicate a true start site and homologies in closely-related organisms.

The occurrence of introns and alternative splicing makes Eukaryotic gene prediction more complex. The prediction is based on recognizing signal functions on



DNA strand and posterior accurate prediction of gene structure and organization. Sequences are classified as coding and non-coding and functions are assigned by alignments with functionally-related documented sequences, involved in transcription, translation and splicing.

After prediction, annotation can be done automatically, by align similarity with sequences at databases and transferring the information about their function to the new sequence. Databases can hold information of nucleic acids (e.g. NCBI, EMBL), protein sequences (e.g. UniProtKB), domains (e.g. Pfam) and structures (e.g. Protein Data Bank), metabolic pathways, gene expression (e.g. dbEST), gene ontology and evolutionary relationships. Different databases are integrated by automatic annotation services, as provided by RAST (Rapid Annotation using Subsystem Technology). Automatic annotation can provide a great amount of information about a new genome from its sequence. However, not all databases are fully curated, updated or have standard terminology. Manual curation is done by checking the results for different databases to identify protein domains and functions, correct start codon positions, gene names, gene products and to identification of frameshift.

After annotation, the genome can be submitted to a database, as submission is a pre-requisite to publication in scientific journals. Once this data is available, it can be used researches about genetic diversity, evolution, biotechnology and health.

### **Comparative Genomics**

After sequencing, assembly and annotation of a genome, different analyzes will try to find out to remove as much information as possible from a genome of an organism. In this context, there is the comparative genomics in order to seek from genomic data, which are the between intra and inter-specific relationships. This can be studied using the homology the genes, loci, functions, processes and categories.

Currently, there are a number of tools available for the comparison of genomes at the level of their genes and proteins, among them: BLAST (Altschul et al., 1997), but their visualization is somewhat limited, hindering an overview of the genomes studied. However, emerged over the last decades other tools, which integrate different databases and integrate visualization and their processing, widely used by bioinformatics are they: CGView (Grant JR, Stothard P, 2008) and BLAST Ring Image Generator (BRIG). With these tools you can see the structure and organization of genes in one or more genomes through a graphical interface (Figure 6). The other programs, such as the PIPS, are

already able to report data pathogenicity, improving the entire information and providing the user various relevant information to do exploration of the body (Soares S et al., 2012).

All of these tools require a computing power to process a large amount of data, since genomes can be processed genomes from of simple organisms (prokaryotes) to more complex organisms (eukaryotes). Another important detail to highlight is the fact that numerous incomplete genomes are being deposited in the National Center for Biotechnology Information (NCBI). Incomplete genomes are those assigned "draft", which will have a percentage of genetic information lost one time analyzed using these tools, this is because part of the "gap" cannot be correlated to anything since it is said to be unknown. Therefore, we emphasize the importance of finalizing a genome in all steps, assembly and annotation, thus avoiding the loss of biological information (Barbosa EGV, 2014).

## **Functional genomics**

### **Transcriptomics**

#### **Concepts and evolution**

Transcriptomics is a global study of gene expression levels of all the genes in an organism (genome-wide expression profiling), in order words, a set of detectably transcripts in a given cell. Transcripts can be modulated diversified factors, such as environment, time, diseases and response to adverse processes.

The first idea of transcriptome emerged in the late 70's, described as the Dot-Blot technique in order to use nucleic acids for the simultaneous analysis of some genes in arrays (Kafatos et al., 1979). In 1995, with the evolution of different technologies, this method has been enhanced by improving their level of sensitivity in optical detection (Brown et. al.). After that, in 2002, the Affymetrix launch the identification of previously unknown genes and detection of splice variants. Finally, in 2008, the RNA-seq technique revolutionized history with sequencing from new generation technologies.

#### **Different RNAs**

Total RNA is the combination of coding and noncoding; being the first comprises the messenger RNA (mRNA), transcript that carries the code information for proteins synthesis. In prokaryotes, the mRNA can be transcribed in operons or single genes and contain the exact transcript, and in eukaryotes, the mRNA has fragments that are translated in proteins (exons) and other regulatory regions (introns) that are removed by a process call splicing.

The noncoding RNAs are a combination of different classes: ribosomal RNA (rRNA), transfer RNA (tRNA), small RNA (sRNA) and long noncoding RNA (lncRNA). This topic will focuses in the non-coding ones, since the coding ones represents less than 2% of the total transcripts of human cells, and to be more precisely the emphases will be given to the regulatory RNAs, which are need to maintain genetic integrity and gene expression in a fine tune. There is still a lot to find out about those types of RNA and as a result of the use of new technologies there set are continually rising.

The tRNAs were known to be responsible for transport specific amino acids for protein synthesis and also recognize their specific codons, but recently another function were attribute to this RNA, when they are exposed to environmental stress in eukaryotes. After cleaved by ribonucleases, tRNAs derived fragments are able to inhibits translation and consequently forms stress granules where the non-translated mRNAs are storages (Kato and Slack, 2013). Therewith, now the tRNAs are also recognized as a new class of regulatory RNAs in response to stress.

The sRNA are responsible for mediate post-transcriptional gene silence (PTGS) and transcriptional gene silence (TGS). In PTGS, the target is cleaved or transcriptionally repress after is bound with the small RNA. TGS is a protective mechanism of genome integrity that keeps a heterochromatic state through DNA methylation or histone modifications (Sabin, Délas and Hannon, 2013a). One of the most known type of sRNA, in eukaryotes, are the miRNA, noncoding RNAs responsible for regulates the expression of genes in a sequencing specific matter. Others types are siRNA, piRNA and sRNAs (prokaryotes). The siRNAs are double-stranded RNA (dsRNA), like miRNA, which directs transcriptional silence of targets, participates in histone modification, heterochromatin formation. Piwi-interacting RNA (piRNA) is a defense system against transposons that acts through deposition of modifications in histones, inducing a heterochromatic state. The piRNAs are present in the genome in clusters and are a germ line specific class from both male and female organisms, and

different from the others classes of sRNAs, piRNA requires the Piwi protein to be generated (Sabin, Délas and Hannon, 2013b).

In prokaryotes, regulatory noncoding RNAs are transcripts of 50 to 250 nucleotides known as sRNAs that can either bound to mRNA and regulates gene expression or bind to proteins and alter their function. sRNAs can act on a target in cis or trans, and are involved in stress response, virulence, quorum sensing, DNA elimination, among others. A recently discover mechanism is the clustered regularly interspaced short palindromic repeat (CRISPR) pathway which are responsible for direct sequence-specific cleavage of foreign DNA (Sabin, Délas and Hannon, 2013b).

Another type of regulatory RNA are the lncRNAs, transcripts with proximally 200 nucleotides or more that is not translate into proteins. They include different classes of RNA like enhancer RNAs (eRNA), small nucleolar RNA (snoRNA), intergenic and overlapping transcripts (Batista and Chang, 2013). Independent of the class, the lncRNA are cell specific with high selective pressure, considering the genetic sequences, the promoters and there function structure, and also can function in cis and in trans, affecting neighbor genes or more distant ones (Fatica and Bozzoni, 2014).

An important aspect of the lncRNAs is that they interact with diverse proteins, forms nucleus domains, interacting two distinct chromosomal regions and modulates the chromatin (Wang and Chang, 2011). The lncRNA can act as a regulatory element, having an influence in histone modification, activating and repressing, even silence alleles of several genes (genomic imprinting) (Lee and Bartolomei, 2013). One example is XIST, a lncRNA involved in the inactivation of the X chromosome in mammals, and aberrant XIST showed to be involved in human cancer and an alteration on a XIST regulator (TSIX – antisense noncoding regulator) result in alteration X inactivation and embryonic lethality (Fatica and Bozzoni, 2014). lncRNA are also involved in the development of central nervous system, heart among other organs and environment responses.

The field of study of the ncRNA is just getting started and there is still a lot to learn about then and what they influence in the eukaryote and prokaryotes life, and because of that the scientific world is focusing all efforts on the study of those transcripts, the real responsible for the complexly of organisms.

## **Methodology of study**

### **Advantages and disadvantages**

## **Microarray X RNA-seq**

Microarray and RNA-seq are experimental techniques to measure the expression levels of the transcripts. These technologies have driven the functional genomics research in different organisms, prokaryotes and eukaryotes, for studies of physiological or pathological conditions. Microarrays allow a quantification of nucleic acids (mRNA form of genomic DNA or cDNA) by hybridization to complementary array. However, the RNA-seq technique, by sequencing of RNAs, can apply the same protocol for various interests, such as total RNA, mRNA, small RNA, such as miRNA, tRNA and rRNA.

There are a number of advantages and disadvantages about these two techniques. Microarray techniques are limited to detect transcripts, whose genomic sequences are their known; however, the microarrays are excellent for specific identification of known variants, for example in diagnostic testing, such as: cystic fibrosis. Nevertheless, RNA-seq is used for detection of known and unknown transcripts sequences and detects splicing and variants, especially you used for further depth studies in any organism, prokaryotes and eukaryotes.

## **Real Time PCR**

Real time PCR is a golden standard technique of gene expression that detects the product as soon it amplifies. The sample (RNA of study) is first reverse transcribed and then a mix similar with the conversional PCR plus a fluorescent agent is added in the sample for amplification. Real time PCR are applied in a variety of biological fields such as microbiology, veterinary science, pharmacology, biotechnology among others and can be useful in detecting SNP variation, allelic discrimination, forensic studies, quantification of pathogens and genetically modified organisms (GMOs), food quality, besides gene expression and ncRNA analysis (Derveaux, Vandesompele and Hellemans, 2010; Navarro et al., 2015).

This methodology has some advantages over the conversional PCR considering specially the measure of the DNA concentration, and another relevant fact is that the results obtained can be either qualitative or quantitative (qPCR), different from de conventional. Real time PCR is widely used technique and so became necessary an experimental protocol with detail information, being a simple and efficient instrument of research is require a maximal quality control so the results can be more reliable and

possible to compare (Navarro et al., 2015). Real-time PCR is very good for the few genes already known, however when do you want to see all a comprehensive analysis already exist other techniques are more suitable and effective.

### **Applied biotechnology: Looking in to the future**

In the past, the focuses of genetics studies were or in the genome or in proteins, but the discovery of new classes of RNA and what they can do opened up a whole world of different approaches to old problems. Besides the knowledge of the biology of the specie at a given moment and situation, obtained especially by the mRNA, now the RNA is also being used in medical treatment, agriculture improvement, and in others fields involving biotechnology.

The development of global transcriptome profile using next generation sequencing as a functional genomics tool allows a better selection of possible target for further studies. With that in mind plus the fact that constant improvement and discoveries are being made in the field of science, the techniques applied now that focused on RNA made incredible breakthrough and are a promise in various fields of research.

### **Ageing**

The use of noncoding RNA like miRNA was first tested in the improvement of lifespan by Boehm and Slack in *Caenorhabditis elegans* and from that different authors pursued this topic (Boehm and Slack., 2005). The miRNAs modulates protein homeostasis, mitochondrial metabolism, stress response, including DNA damage (Kato and Slack, 2013). Using new sequencing technologies among other, researchers discovered that many *C. elegans* miRNA including *lin-4* changed their expression during life and in the absence of *lin-4* the animals decreased their lifespan (Kato *et al.* 2011). This and other breakthroughs demonstrated that miRNAs as capable of predicting the longevity and remaining lifespan, and that possibly others ncRNAs are also involved in ageing (Kato and Slack, 2013).

### **Gene function and vaccine production**

Next generation sequencing became an extremely important tool for the understanding of gene function, since the costs of sequencing for the whole genome and transcriptome is more cheap and fast than previously methodologies. The study of the transcriptome allows us to understand what happens with the microorganism and the

host during the infection and select potential virulent factors with more precision. This selection is possible thanks to studies of different expression in environment (e.g. culture medium) that simulates the conditions faced by the pathogen or by a double transcriptome study where both the microorganism and the host are analyzed *in vitro* or *in vivo*.

### **Insect management**

Insect pests are a major problem in agriculture and the use of pesticides are slowly decreasing since it is toxic for humans, fact that brought the necessity for nontoxic approaches. RNA interference (RNAi), a process of specific gene silencing based on dsRNA, was successfully tested in deleting essential genes in insects and by doing so caused lethality (Baum et al., 2007). Both miRNA and siRNA are pathways endogenous of insects and plants, and the siRNA are recognized as a major antiviral defense mechanism of insects. Even so this strategy proved to have potential not only in killing pests but also in the control of diseases transmitted by insects (Burand and Hunter 2013). The problem encountered for this method was the way of delivery, since dsRNA does not replicate in the host and because of that need to be constantly supplied. Two popular approaches are the use of RNAi pesticides that are safe for humans and beneficial insects, and use of transgenic plants that already produce dsRNAs (Burand and Hunter, 2013) but in this case brings further problems that will not be discussed here.

### **Disease therapy**

The RNAi therapy shown to be effective in treating a variety of disease from viral disease to single nucleotide polymorphisms (SNP) and oncogenes, but as mentioned in the topic above the delivery and possible toxicity of the RNAi are still a concern (Hannon and Rossi, 2004).

Human immunodeficiency virus (HIV) was one of the first targets by RNAi in the attempt to silence an infectiousness target, but since it is a highly mutated virus there is still a lot to figure it out until this approach be successful. A second method was then tested to improve the effectiveness of the treatment which instead of focusing on the virus the target was a host receptor required for the infection (Boden et al., 2003). The use of RNAi to treat disease are in progress and apart from HIV other illnesses are also in experimental testing and among them are hepatitis C virus (HCV), amyotrophic lateral

sclerosis (ALS), Huntington, different types of cancer and more (Hannon and Rossi, 2004; Havens, Duelli and Hastings, 2013).

Approaches that manipulates splicing also function as disease therapy by correcting aberrant splicing or inducing then. Mutation that disrupt or create new splice sites are very common, in both intronic and exonic location, and are known to cause genetic diseases (e.g. Duchenne muscular dystrophy, Miyoshi myopathy, Cancer) (Havens, Duelli and Hastings, 2013).

## **Proteomics**

The development of new technologies, involved in next-generation sequencing technology and access the complete genomes has promoted an increase of information about genomic data for several organisms. Thus, the functional genomics come with aim of complementing genomic data, where through of transcriptomic and proteomic studies, has promoted the description of the gene and protein functions as well their interactions. However, unlike of genomic and transcriptomic studies, the proteomic try to evaluate the protein expression at a given condition. Thus, a proteomic study provides valuable information about: protein synthesis, qualitative and quantitative information of protein product, post-translational modifications, protein-protein interaction and subcellular localization.

Currently, due the great advances both in classical and high-throughput proteomic technologies, new advances has been observed in the identification of targets for infectious diseases, inflammatory diseases, chronic conditions, cancer, drug discovery. Furthermore, through of comparative proteomic screening, between two conditions, i.e normal and diseased cells, allow the identification of differentially expressed proteins, that can promoted the identification of novel targets and biomarkers (Heffner et al., 2014; Wu et al., 2008).

The proteomic allows the analysis of the gene expression on a large scale; this study is basically divided into two steps: (i) separation of the protein extract through of the intrinsic fundamental properties of each protein and (ii) identification of proteins through of the sequencing of peptides predigested enzymatically. However, the success of a proteomic study is due the complimentarily by bioinformatics analysis; where several softwares are developed to auxiliary in the data analysis, i.e., correlation with genomic data, furthermore, several public data repository for proteomics data were



created. Thus, proteomics is presented as a multidisciplinary discipline in the investigation medical, biologic and biochemistry.

### **Gel-based proteomic**

Gel-based methods like gel electrophoresis are routinely used in proteomic analysis to separation of molecules or peptides. Two-dimensional gel electrophoresis (2-DE) is a classical proteome technique, applied to resolution of protein from complex protein mixture. 2-DE is used to separate proteins in two dimensions: (i) during the process of isoelectric focalization (IEF) the proteins are separated according to the isoelectric point (pI) and (ii) using a SDS-PAGE the proteins are separate by molecular weight. After the electrophoretic resolution the spot/protein are submitted to tryptic digestion and identified by mass spectrometry (MS) (Görg et al., 2004). Thus, the combination of the 2-DE with the technique MALDI-TOF-MS (matrix-assisted laser desorption/ionization – time of flight), has promoted great insight in the study proteomics.

The 2-DE analysis allows the generation of proteins maps, where spots/proteins differentially expressed are visualized, as well the presence of post-translational modifications (PTM) and isoform of proteins. However, a disadvantage of 2-DE classical is gel-to-gel variation, thus to suppress this variation the two-dimensional fluorescence difference gel electrophoresis (2-D DIGE) was developed (Unlü et al., 1997). This technique allows the simultaneous separation of two samples different pre-labeled with cyanine dye (CyDye) fluorophores in the same 2-DE. Furthermore, the great advance of this technique is the realization of a multiplex assay, where an internal standard is used for normalization of data that minimize the experimental variation between gels and favor the quantitative analysis of the proteins (Arentz *et al.* 2014).

### **Gel-free proteomic**

Due the great advances in the biotechnological, new technologies also has been developed in proteomic area, where several advances have been observed in recent years, in order to address the weaknesses observed in 2-DE. Thus, the development of gel-free methods, such as liquid chromatography (LC) MS/MS promoted a great advance in the proteomic field, where are obtained higher sensitivity, accuracy and resolution in proteomic study of large-scale. In this approach normally utilize a microcapillary reverse phase (RP)  $\mu$ LC system to separate the protein extract previously

cleaved into peptides and *in tandem* analyzed by MS/MS, using electrospray ionization (ESI) coupled the mass analyzer (time-of-flight (TOF), quadrupole and ion trap). Thus, LC-MS/MS analysis associated the database searching has promoted a high-throughput proteomic analysis (Roe and Griffin, 2006).

The abundance changes of proteins present in complex biological mixtures is a major challenge in the proteomic study, thus with the introduction of gel-free methods, several workflows to quantification have been developed such as: (1) Chemical labeling: utilize thiol-specific tags to covalently label in cysteine residues through of the isotope-coded affinity tag (ICAT) reagent approach, in addition other approach used to quantification using the chemical labeling is the technique iTRAQ, where is based in the labeling of an amine reactive, isobaric, isotope tag. (2) Metabolic labeling: metabolic precursors (amino acids, carbohydrates, organic salts) are introduces in the culture media and changed the molecular weight of the proteins, these metabolic precursors are labeling, through of the technique of stable isotope labeling of amino acids in cell culture (SILAC). (3) Label-free quantification: no label type is used to quantify the samples. The quantification is based on (i) spectral counting; where the quantification is achieved by total number of MS/MS spectra for a protein, and (ii) intensity of the chromatographic peak area under the curve or intensity of the precursor ion MS spectra (Otto, Becher and Schmidt, 2014).

### **Proteomic in applied microbiology and biotechnology**

The proteomic study offer a global survey of cellular changes in different physiological states at protein level, thus several studies using distinct or combined proteomic approach has been applied both in the applied microbiology and biotechnology.

In the microbiology field several proteomic studies have been performed with lactic bacteria, due the importance of this group of bacteria in the aliment industry, act as components of probiotic products and mainly in the fermentation process of yoghurt and cheese as well in the long-term storage of these products. Thus, proteomic studies related to knowledge of the process adaptive of *Lactobacillus*, *Propionibacterium* and *Bifidobacterium*, to acid condition and bile salts, showed a set of proteins that favor the survival and adapt theses bacteria in theses environmental that are present in the gastrointestinal tract (Leverrier et al., 2003; Sanchez et al., 2007; Hamon et al., 2012). In addition, the combination of bioinformatics tools and proteomic techniques has been

applied to explore the host–bacteria interaction process, through of the characterization of surface-exposed proteins main protein fraction used by bacteria to promote the interaction with receptors of the host-cell (Barinov et al., 2009; Meyrand et al., 2013). However, these studies favor the selection of strain bacterial that can be used in the alimentary industry.

The genomic data from cancer patients has promoted information about variations, mutations and molecular pathways this disease. However, these alterations reflect in the gene expression of the cell cancer and consequently affected the function and stability of the proteins coding by these genes. Thus, proteomic study allow unveil this alteration and identify proteins involved in this pathway. In addition, currently the proteomic has been applied in the research of potential biomarker to have utilized in diagnostic of various type of cancer. In this type of study the biological samples evaluated from blood specimens like serum or plasma, thus quantitative proteomic both label-based and label-free has promoted the identification of several regulated differentially proteins, when compared the normal cell and tumor cell (Barinov et al., 2009; Kocevar, Hudler and Komel, 2013; Meyrand et al., 2013).

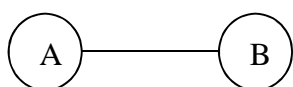
Currently with the technology advances, great efforts have been directed to the study of sustainable energy using mainly the bioenergetics process, which aims at energy production from biological materials and photosynthetic organism. A of the applied bioenergetics process is at production of biofuels. Thus, several proteomic studies are used to characterize the main raw material used to produce biofuels. Proteomic study performed with the plants: Sorghum sugarcane and maize, and algae and cyanobacteria that also are used in biofuels fabrication, have promoted the characterization of proteins present in different subcellular compartments, involved in several biological process as well the identification of proteins associated the resistance and adaptation to several stress conditions. These studies combined different proteomic approach and the results obtained in these studies theses several source used in the bio-ethanol production, promote information about of the physiology of this organism as well the identification of targets that can be used in its improvement main during the cultivation in different climatic conditions, favoring the biofuels production (Ndimba et al., 2013).

## **Application to a bacterial protein-protein interaction**

With the advent of second and third generation of sequencing technologies of there has a significant reduction of cost and time, and simultaneously happend an amount of sequenced genomes. With that, the genome sequencing became more accessible the genomes sequencing, therefore increasing the amount of biological data such as DNA, RNA and proteins available in public databases about of various organisms (Casals et al., 2012; Koboldt et al., 2013; van Dijk et al., 2014a).

DNA, RNA and proteins do not act individually in a living organism they work together interacting with each other other, physicaly or regulatory, in such a way they can execute their biological processes. Knowing whereby these elements are grouped and how they interact is a relevant area of study and research, allowing the understanding of a given organism, disease or biological process at systemic level. Therefore, the data generated from DNA, RNA, proteins and other molecules, permit researchers to broaden the fields of research to better understand organisms and/or diseases, search for new pharmaceuticals, veterinary, agricultural and biotechnology compounds.

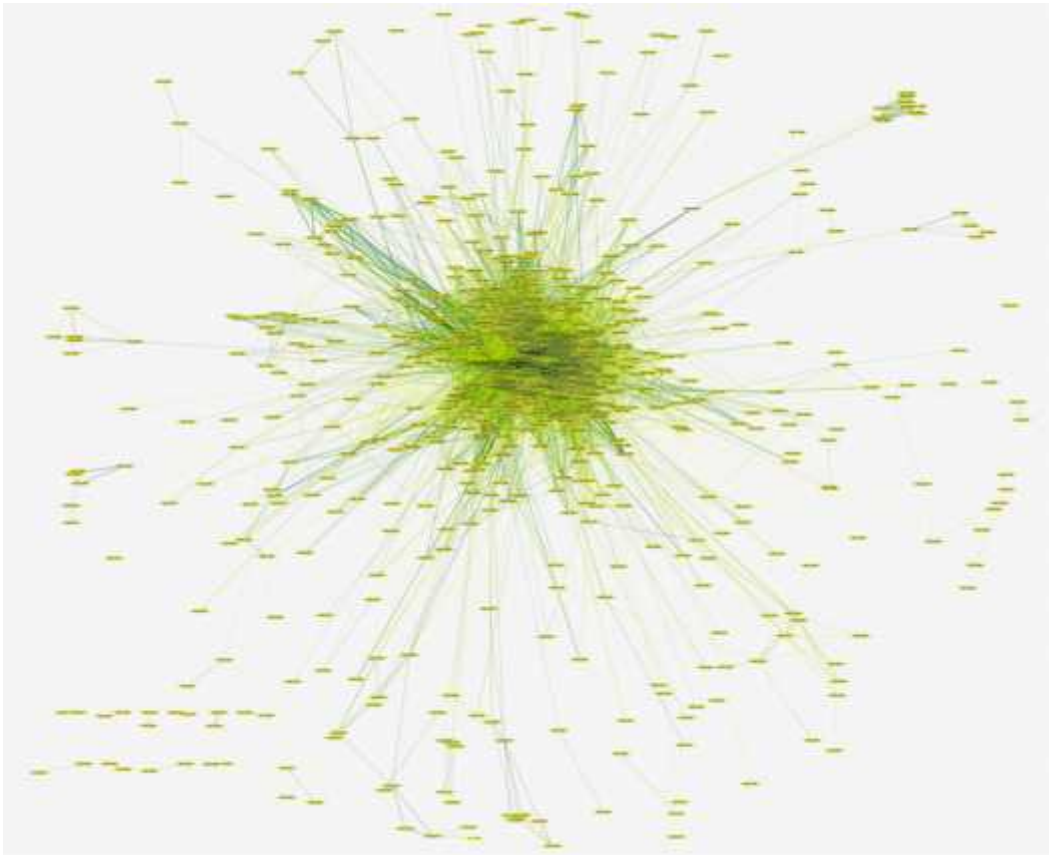
An important tool for understanding how these molecules act in complex biological systems are protein-protein interactions (PPI) that, pair-to-pair, form a complex set of biological interactions called interactome (Ngounou Wetie et al., 2014; Rao et al., 2014). A PPI is composed of nodes that represent proteins (A and B), and edges that connect these proteins (nodes). The edge that connects both proteins do not have source and destination, thus forming a non-directional interaction, being equivalent the interpretation that the protein interacts with protein B or that protein B interacts with the protein A (Figure 5).



**Figure 5: Representation of the interaction between two proteins using the graphs theory where the nodes A and B are connected by an edge.**

Each interaction has two elements, called binary, and several binary interactions form a complex network of PPI that may represent all interactions of an organism or pathogen-host (Figure 6). The literature describes methods for the identification of PPI, being classified as: (i) experimental - yeast-two-hybrid (Y2H), protein chip, tandem affinity purification followed by mass spectrometry (TAP-MS); - biophysical - atomic force microscopy (AFM), analytical ultracentrifugation (UC), nuclear magnetic

resonance (NMR) or (ii) computational - three-dimensional structure, text mining, phylogenetic profile, phylogenetic tree, gene co-localization, interolog mapping-based and machine learning.



**Figure 6: Example of a protein-protein interaction network, made up of several pairs of interaction.**

In time, the methods can be classified by data generated as low-throughput or high-throughput. High-throughput methods are able to detect greater amount of interactions, but have low specificity, while the low-throughput methods, which test specific interactions, have high specificity. More specific is the method, most sure has that the interaction occurs within the organism and, the more sensitive is the method, the more interactions are known. Experimental and computational methods complement each other, both being necessary for the scientific community to better explore the data (Harrington, Jensen and Bork, 2008; Mulder et al., 2014).

Regardless of the method used, a network interaction can provide researchers with more knowledge about an organism, making possible to think about new biological hypotheses and direct new experiments. In this sense, though indirectly, network interaction can be used for biotechnological purposes. By better knowing the organism, this can be genetically modified and have his behavior changed, aiming at greater

efficiency or quality in the production of a product of our interest, as well as, can be genetically modified to produce or metabolize a product that is harmful to human, animal health or the environment, such as bio-remediation of soils. Aided by a network of PPI this is possible.

Assuming that some gene, óperon or protein involved in a biotechnological process of interest be known, with the help of a PPI network, it is possible to identify the proteins that interact with (partner) the protein of interest. By identifying proteins an interaction, may also be known their functions allowing to better understand the biological process or even the biochemical pathway involved in biotechnological process of interest. This information, aggregated to the knowledge about the organism, can help a researcher to understand the organism at systemic level, allowing new hypotheses to be tested in the laboratory. Still, with a network of interaction topological analysis can performed to identify biologically important proteins as proteins Hubs that have a high degree of interaction with other proteins, and proteins Betweenness that make the connection between the functional modules in a network of interaction. These both protein class, Hub and Betweenness, are considered essential proteins (Mulder et al., 2014).

Information of interactions, protein Hubs and betweenness allies proteomic experiments or information to RNA-Seq, which test a condition of biotechnological interest and measure the expression of transcripts or proteins, extend the prospects of identifying targets or metabolic pathways relevant. Besides allowing know which existing proteins and the expression of these, it is possible to meet with who these proteins interact and work together for the realization of biotechnological condition tested.

**Anexo 2: Capítulo de livro: *RNA-seq: Reveling biological insights in Bacteria* aceito em *Next Generation Sequencing* (disponível em breve em InTech - Open Access Publisher).**

**Annexe 2 : Chapitre du livre : *RNA-seq: Reveling biological insights in Bacteria* accepté dans *Next Generation Sequencing* (bientôt disponible sur InTech - Open Access Publisher).**

MARIANA P SANTANA<sup>1</sup>; FLAVIA F ABURJAILE<sup>1</sup>; MARIANA T D PARISE<sup>1</sup>; SANDEEP TIWARI<sup>1</sup>; ARTUR SILVA<sup>2</sup>; VASCO AZEVEDO<sup>1\*</sup>; ANNE CYBELE PINTO<sup>1</sup>

\*Corresponding Author: Vasco Azevedo

E-mail: vasco@icb.ufmg.br, Telephone +55(31)34092610 Fax: +55(31)34092610

<sup>1</sup>Instituto de Ciências Biológicas-ICB/UFMG. Departamento de Biologia Geral. Avenida Antônio Carlos, 6627, Pampulha. Belo Horizonte, Minas Gerais, Brasil. Caixa-postal: 486/CEP: 31270-901

<sup>2</sup>Centro de Ciências Biológicas, Departamento de Genética. Universidade Federal do Pará, Campus do Guamá, Guamá. Belém, Pará, Brasil. Caixa-postal: 8607/CEP: 66075-900

**Abstract:** New technologies are being constantly released and those improvements carry by then bring advance to not only on transcriptome, the focus of this book chapter, but also to diverse areas of biological research. Since the announcement and application of the RNA-seq approach, discoveries are being made in this field and when we consider bacterial species, this progress proceeded a few years behind. However, with the application of RNA-seq derivatives approaches, we can reveal biologicals insights in Bacteria world and thus aspire to uncover the mysteries involving gene expression, organization and others functional genomics features.

**Keywords:** RNA-seq, bacteria, transcriptomics, bioinformatic analysis workflow.

## **1. Introduction**

The RNA-seq technology has driven advances in gene expression analysis, through the new generation sequencing platforms, being versatile, powerful, ensuring quality results with accuracy and reproducibility never reached before. This technology generates information that designates significance to the set of transcripts (transcriptome), opening up possibilities to understand cell behavior in different environments. RNA is a component of excellence within the cell, since it plays different roles as messenger regulatory molecule, carrier and is essential for maintenance of housekeeping genes (Creecy and Conway 2015).

In 2005, the first new generation of sequencing technology was released and has since been evolving rapidly (Margulies et al. 2005). After starting the process of gene expression analysis in bacteria (Passalacqua et al. 2009; Perkins et al. 2009), which a

more accessible cost, shorter time experiment and without probe, the technology took off and today overlaps other tools used for this purpose, such as the microarray technology, till now extremely used for this type of analysis.

## **2. Applications of RNA-seq**

The understanding of the transcriptome is essential to the knowledge of functional genomics of an organism. The development of next-generation sequencing (NGS) impact different areas, such as medical and industrial, and have gone through a revolutionary process where different approaches have emerged in the field of microbiology and molecular biology, among them the RNA-seq technique, in order to aid understand and bring answers to the bacterial domain investigations. In this section, we will detail some of these applications that are part of our current context.

### **2.1. The medical field**

The implication of these NGS technologies in medicine allowed the expansion in the field of diagnostic, treatment and prevention, especially concerning the understanding of bacterial diseases. One of their major applications have been quantification of expression levels of each transcript under different conditions that simulate the intracellular environment, such as the work of Pinto and collaborates (2014) in order to understand the relationship between host-pathogen (Pinto et al. 2014). The validation of this technique to the transcriptome of the pathogenic bacteria as their host was demonstrated in the work of Westermann and collaborates (2012) through the dual RNA-seq, that simultaneously analysed gene expression of the pathogen and host (Westermann et al. 2012). This allows us to better understand the systems biology between bacteria and their host, helping scientists to develop drugs and vaccines.

Another field that has been much explored is the research involving metatranscriptome, as scientists seek to comprehend the composition and regulation of microbial ecosystems (Leimena et al. 2013; Maklaim et al. 2013) and pursuing this, they use the RNA-seq technique to generate a large volume of data what grants large reliability and allows a broad interpretation of the data. The study Leimena and collaborates (2013) validated RNA-seq technique in the microbiota of human's small intestine with ileostomy, aiming to understand the interactions involved in this microbial ecosystem and how these relationships can be associated with diseases (Leimena et al. 2013). Transcriptome analysis pipelines (see topic 5) can be used for different experimental designs, applied for many bacteria, besides those on the medical field.

### **2.2. The industrial field**

Industrial applications have been developed over the past few years, mainly around the probiotic industry, since it brings gains to the world economy. The work Bisanz and collaborates (2014) shows metatranscriptome of probiotic yogurt through RNA-seq technique (Bisanz et al. 2014), seeking the understanding of metabolic activities that allows the survival of this organism in the products. The results show the adaptive capacity of this bacterium as well as the variation in differential gene expression according to the taste or storage life of the product (Bisanz et al. 2014).



Studies such as these are of great importance, because enrich the knowledge of the industrial field, opening new perspectives for an attractive area in the marketplace, which results in an improvement in the quality of the final product to be delivered to the consumer.

In addition to the probiotic market, another major importance sector is the production and synthesis of biomolecules by bacteria. Studies such as Wiegand and collaborates (2013) sought to understand through the RNA-seq technique the regulatory RNAs during fermentation of *Bacillus licheniformis*. This study identified active genomic regions, which in turn contribute to the efficiency and optimization of the fermentation process, which can promote the industrial production of exoenzyme and antibiotics (Wiegand et al. 2013).

Microorganisms produce antioxidant molecules can be used in the pharmaceutical and cosmetic industry, as well as other compounds, such as propionate, applicable in the production of chemicals aids and produced by *Propionibacterium freudenreichii ssp. shermanii*, considered valuable on the Food Industry (Wang and Yang 2013b). In this scope, the RNA-seq technology is very promising and its application is able to bring advances polls in these studies.

### **3. RNA-seq and derivative techniques**

#### **3.1. RNA-seq**

The RNA-seq technology is able to directly and quantitatively identify all RNAs: coding and non-coding; rare and abundant; smaller and larger. This method provides the transcription start site information (TSS), untranslated region (UTRs), detection of unknown open reading frames (ORFs), improving quality in genomic annotation (Sorek and Cossart 2010), and also allows the distinction between primary transcript and processed (dRNA-seq) (Sharma et al. 2010).

The major constraint is to ensure representatives for rare transcripts, and in this case, the recommendation is to increase the representation of reads per library (Haas et al. 2012) or enhance these transcripts eliminating the ribosomal (rRNA) and transfer (tRNA), that are in abundance in the cell representing about 95% of total RNA (Bischeler et al. 2015).

Despite being considered the gold standard in gene expression analysis, for some researchers is complicated to define this technology as the gold standard, a method that presents itself in different platforms and address different strategies, showing advantages and disadvantages. However, they do not question the superiority of the technology compared to other in the past (Kratz and Carnin 2014).

Even with all superiority, we cannot rule out the need for biological replicates and depth of sequencing, hence the results may achieve greater reliability and reproducibility (Sandler, Johnson and Krawetz 2011). Differentially expressed genes are better appraised when we have samples with more biological replicates compared to an enhanced depth with less replicas (Liu, Zhou and White 2014).

Transcriptomics studies have provided a revolution in the bacterial environment. Different bacterial species have been targeted for RNA-seq studies (Sharma et al. 2010; Isabella and Clark 2011; Pinto et al. 2014; Patenge et al. 2015) and gene expression-

based discovery has transformed the scientific paradigm of these organisms. The detection of an unexpected amount of coding genes in *Helicobacter pylori* demonstrated that despite having a small compact genome the transcriptome of this bacteria is extremely complex (Sharma et al. 2010).

A surprising result was the detection of a large number of Transcription Start Sites (TSS) never achieved before by any technology, aside from derivative RNA-seq technology like the differential RNA-seq (dRNA-seq), which was able to differentiate primary transcripts that exhibit triphosphate end from processed transcripts that presents monophosphate end as rRNAs and tRNAs. In this case, to enrich mRNA, the strategy was the treatment of total RNA samples with exonuclease enzymes that degrade nucleotide monophosphate. This strategy identified 5'UTR end, operons, antisense transcription, revealing a new perception in the organization of the bacterial transcriptome and a new vision for analysis of individual genes (Sharma et al. 2010).

These results obtained allows us to infer a role of 5'UTR regions, and a correlation between size and cell function was proposed by the researches, whom found that larger size is related to pathogenicity (Sharma et al. 2010). These results show how little knowledge we have regarding microorganisms, believed to be the simplest form of life, and yet proves to be more complex than previously anticipated which still a lot to discover.

A RNA-seq application that has been widely used in bacterial genomes are studies focused to identify smalls RNAs (sRNA). These elements are regulators of various biological processes and were initially and mostly studied in *Escherichia coli* (Gottesman and Storz 2011). However, with the advances in technology, it has been possible to identify and characterize small RNA in a variety of bacterial species (Papenfert and Vogel 2010; Sharma et al. 2010; Kroger et al. 2012) . Yan and collaborates (2013) identified in *Yersinia pestis* an expression profile of sRNA in vitro and in vivo, which allowed the identification of new sRNAs and the recognition of gene expression modulation during the infection process, improving the understanding of transcription regulation mechanisms of this organism (Yan et al. 2013). The importance of these studies involving sRNA also includes assistance in research related with antibiotics therapies, a study in initial development albeit which a lot of knowledge to be better exploited (Storz, Vogel and Wassarman 2011).

RNA-seq has been used in different areas and situations, advanced studies with the technology can detect details in the cell expression (van Dijk et al. 2014b). Even with the difficulties in separating eukaryotic and prokaryotic material, it was possible to distinguish the simultaneous expression profile between the host-pathogen responses through dual transcriptome studies. This work allowed unknots the host response against the bacterial infection and virulence factors enabled to determine the infectious process (Humphrys et al. 2013). These studies contribute to the research in the field of biological infection, considering diverse pathogens with different life cycles and methods of infection, granting crucial knowledge for studies of diagnostics and vaccines, as metatranscriptomics study.

After a relatively short time in the market, RNA-seq can accurately reveal structural and functional elements of the bacteria. The mapping of transcripts in the

genome can refine the annotation or even identify new regions, improving the quality of the studied genome, compared to regions previously annotated by predictors or assembled by an ab initio approach (Martin and Wang 2011; Tjaden 2015), and can even check out the abundance of transcript expression.

Data arising from a quality genome tends to collaborate with more promising results, answering the biological question being analysed by researchers. In search of a genome quality, ab initio transcripts assembly or even a hybrid approach, which uses both the reference genome as a reference assembly, become an auspicious endeavour to solve many problems encountered in the genome and complicated to adjust (Martin and Wang 2011).

Pinto et al. (2012) conducted a study in *Corynebacterium pseudotuberculosis* adopting ab initio assembly, and therefore was possible to identify differences in the expression of active genes in different environmental conditions, allowing the detection of possible new virulence factors involved in pathogenicity, making it reportedly targets for vaccine development, diagnostic or treatment against caseous lymphadenitis disease caused by this bacterium (Pinto et al. 2012).

These results infer the importance of this technology and the possibility of going further with a tool that aims to improve and probably will expand the field of analysis, making the results increasingly closer to the bacterial molecular reality.

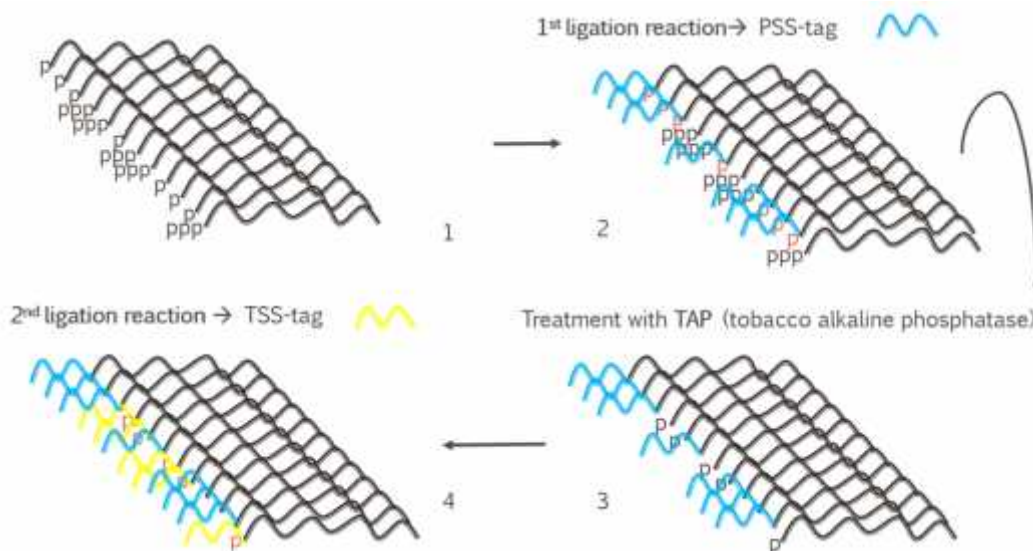
### 3.2. tagRNA-seq

Bacterial RNA can be divided in two groups: primary or processed transcripts. Primary transcripts are represented by the presence of 5' tri-phosphate (5'PPP) which includes messengers RNA (mRNA) and small RNAs (sRNA), and processed transcripts are those carrying 5' mono-phosphate (5'P) such as mature ribosomal RNA (rRNA) and transfer RNA (tRNA).

Transcriptome studies usually focus on mRNAs or sRNAs and because of that, depletion or enrichment reactions are needed to eliminate specially the rRNAs, since they represent proximally 95% of the total bacterial transcriptome (Bischler et al. 2015). A recently developed approach called dRNA-seq (Sharma et al. 2010) revolutionized the research of the primary transcripts by considering the 5' difference between the primary and processed groups, as mentioned previously (see topic 3.1).

RNAs are very stable and during the whole preparation, considering the “whet-lab” experiments, some transcripts are partially or totally degraded. 5'PPP or 5'P are one of the mechanisms of protection against exonucleases and the first degraded portion. During that process, information is lost and some primary transcripts ended up with 5'P and so being treated akin as a processed transcript and consequently being eliminated during the dRNA-seq technique. Considering this flaw, a new methodology was created to overcome this problem by tagging and clustering the two groups together in a RNA-seq derived approach named tagRNA-seq (Innocenti et al. 2014). This technique also considers the difference between processed and primary transcripts, but instead of degrade the processed ones, two different ligation reactions are implemented with two different markers (PSS-tag and TSS-tag). Figure 1 exhibits briefly the methodology considering the three main steps: (1) the first reaction tags (PSS-tag) on

the processed transcripts; (2) treatment with tobacco alkaline phosphatase (TAP), where the 5'PPP loses two phosphate, which allow the third step; (3) the second ligation reaction (TSS-tag) on the primary transcripts. After those steps, the transcripts are sequenced and due to the different markers, they can be distinct and compared (Innocenti et al. 2014).



**Figure 1.** The three main steps of the tagRNA-seq approach. (1) The first ligation reaction, during which the attachment of the PSS-tag (blue) to the processed transcripts (5'P); (2) Treatment with tobacco alkaline phosphatase (TAP), turning triphosphate to monophosphate groups; (3) The second ligation, corresponding to the TSS-tag (yellow) marker on the previously 5'PPP group (primary transcripts). The different markers allow the differentiation of the triphosphate and monophosphate groups after sequencing.

This methodology was first described in *Enterococcus faecalis* (Innocenti et al. 2014) and was based in other technique 5'tagRACE (Fouquier d'Herouel et al. 2011), a 5'RACE derived method. The results provided by tagRNA-seq improved the annotation of the *E. faecalis* genome for having identified or correct several genome portions, including non-coding and coding regions. This study also compared different libraries to prove the effectiveness of this innovative approach and with this provided a new method capable to differentiate primary and processed RNAs, and suited to better comprehend the genetic information of bacteria as other groups (Innocenti et al. 2014).

dRNA-seq and tagRNA-seq are approaches that allow a new view of the transcriptome by selecting the primary transcripts for sequencing or differentiate the primaries from the processed transcripts for a larger insight in the transcriptome. Those state-of-the-art techniques promises a better understanding of RNA structures like TSS, 5'UTR, promoters among others besides the knowledge of non-annotated genes and small RNAs.

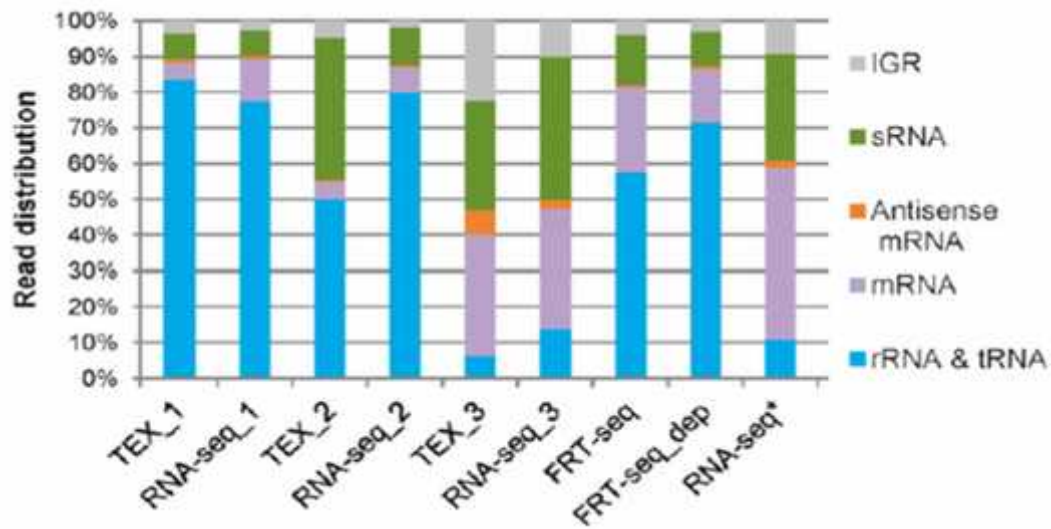
### 3.3. FRT-seq (Flowcell Reverse Transcription Sequencing)

Flowcell reverse transcription sequencing (FRT-seq) is a new and improved methodology derived from the RNA-seq technology that was created for Illumina sequencers. Unlike RNA-seq, FRT-seq does not require amplification by PCR, a step that usually introduces bias in the results by displaying an erroneous view of the quantity of some RNA species (Mamanova and Turner 2011). Another important feature of this Illumina sequencing methodology is the ability to generate strand specific information, use of pair-end library and the need of a considerable initial amount of RNA template. The PCR free amplification is a major step forward a more comprehensive library, akin the original one, and also without formation of intermolecular priming artefacts, among others errors, and probably will be a fairly used technique in a near future (Mamanova et al. 2010; Mamanova and Turner 2011). The third generation sequencing platforms like Nanopore and PacBio also uses amplification free approaches but neither are currently being broadly used since they still exhibits sequencing errors.

The FRT-seq technique comprises the fragmentation of the template (e.g. mRNA) followed by ligation of adapters in both 3'end and 5'end, responsible for the hybridization of the template with oligonucleotides on the flowcell surface. The next steps performed are the quantification, the reverse transcription and then the sequence reaction (Mamanova et al. 2010; Mamanova and Turner 2011).

This approach can be applied for booth eukaryotes and prokaryotes although the amount of published papers involving eukaryotes are more substantial. Considering the Bacteria world, we can quote papers involving *Salmonella enterica* (Kroger et al. 2012) and *Shigella fleneri* (Vergara-Irigaray et al. 2014). The FRT-seq methodology was practiced as a complementary approach to describe the transcriptional landscape of the species, and in both cases, FRT-seq showed greater sensitivity and an excellent concordance when compared to others approaches and replicates.

The *S. enterica* paper showed that FRT-seq is as efficient as any other sequencing methodology, such as RNA-seq and dRNA-seq techniques (Figure 2) (Table 1) (Kroger et al. 2012). Figure 2 compares 9 different RNA libraries (TEX (1, 2, 3); RNA-seq (1, 2, 3, \*); FRT-seq (depleted and not depleted). TEX is a dRNA-seq methodology (see topic 3.1 and 3.2) and together with the first three RNA-seq biological replicates were sequenced in a 454 (1 and 2) or a Illumina GAII (3) sequencer. The charts relates the percentages of different RNA species and shows that the FRT-seq libraries show similar results or even better than the others approaches. The data presented on Table 2 also supports this claim, especially considering the total number of reads and the uniquely mapped reads achieved by FRT-seq libraries.



**Figure 2.** Sequencing methodology comparison. Adapted from Kroger et al. 2012. IGR – Intergenic region; TEX – libraries treated with terminator exonuclease; RNA-seq\* – library enriched for small RNA species (sRNA).

**Table 1.** Sequencing statistics. Adapted from Kroger et al 2012.

Library	Sequencing technology	Description	Total number	Number of reads (not mapped)	Number of reads (uniquely mapped)	Percent uniquely mapped reads [%]	Minimum fold coverage <sup>#</sup>
TEX_1	454	dRNA-seq library biological replicate 1	161,031	72,623	88,408	54.90	1.11
RNA-seq_1	454	RNA-seq library biological replicate 1	248,993	83,030	165,963	66.65	2.03
TEX_2	454	dRNA-seq library biological replicate 2	111,462	10,785	100,677	90.32	2.16
RNA-seq_2	454	RNA-seq library biological replicate 2	93,337	38,577	54,760	58.67	0.61
TEX_3	Illumina GAI	dRNA-seq library biological replicate 3	1,738,867	122,058	1,211,426	69.67	20.99
RNA-seq_3	Illumina GAI	RNA-seq library biological replicate 3	2,148,563	136,871	1,360,113	63.30	21.16
RNA-seq*	Illumina HiSeq	RNA-seq library biological replicate 4	3,750,797	164,658	2,596,010	69.21	25.11
FRT-seq	Illumina GAI	FRT-seq library biological replicate 5	18,563,218	4,203,715	2,456,792	13.23	16.42
FRT-seq dep	Illumina GAI	FRT-seq library biological replicate 5 rRNA depleted	24,585,564	9,652,397	4,093,744	16.65	27.77

The *S. fleneri* paper (Vergara-Irigaray *et al.* 2014) also obtained a favorable result concerning FRT-seq, in fact this approach attained a bigger gene repertoire than the RNA-seq (Table 2).

**Table 2.** Sequencing statistics. Adapted from Innocenti *et al* 2014.

	RNA-seq		FRT-seq	
	Condition A	Condition B	Condition A	Condition B
Total number of mapped reads	20,099,597	22,736,494	49,925,286	47,605,241
Total number of reads mapping to genes	1,525,782	2,271,423	3,037,954	2,585,600
Reads mapping genes in sense	1,195,446	1,958,533	2,469,828	2,129,951
Reads mapping genes in antisense	330,336	312,890	568,126	455,649

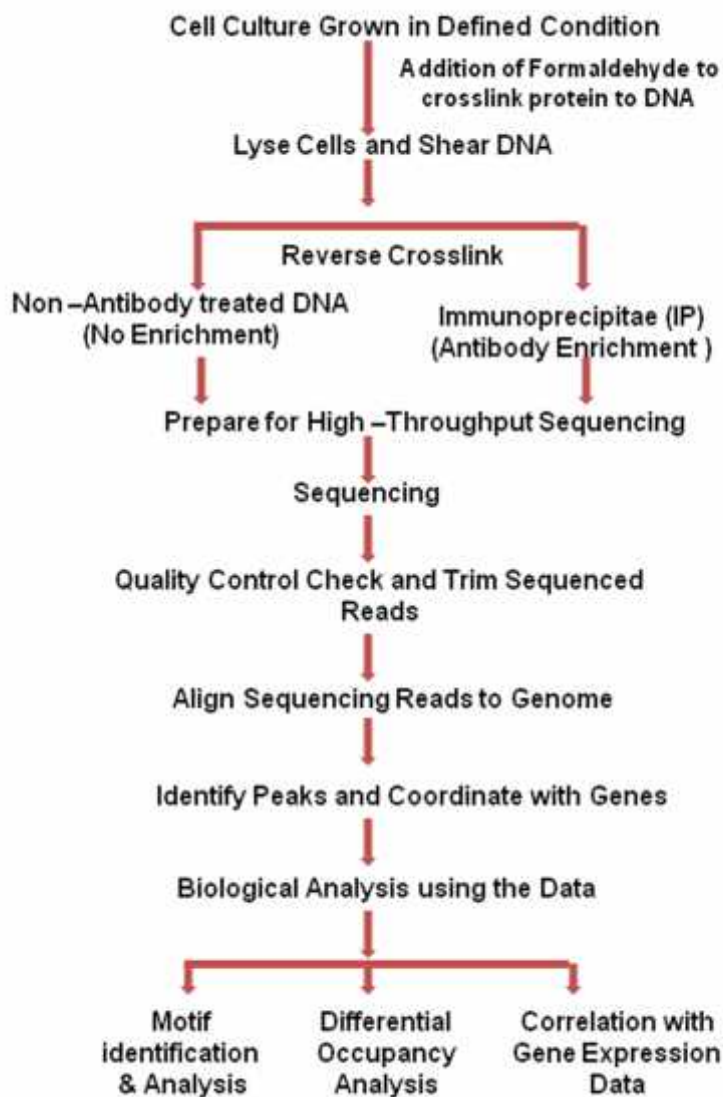
The data provided in this topic demonstrate the quality of this recent publish methodology and according to the authors (Mamanova *et al.* 2010; Mamanova and Turner 2011) news updates are still under improvements what will probably set up an even better approach for the users. The fact that this technique is only applicable for Illumina sequencers is a drawback, but since this sequencing platform is present worldwide, this disadvantage can be simply fixed or perhaps in a near future can be settled to work in others sequencing platform. Another particularity of this technique is the efficiency towards AT-rich genomes what do not constrain the application on AT-poor genomes, this fact is due to the PCR-free amplification what compel also a question on others sequencer like Nanopore and PacBio. Despite these issues, this technology has a bright future and is a great advancement for the conventional RNA-seq.

### 3.4. Chromatin immunoprecipitation followed by sequencing (ChIP- seq)

Chromatin immunoprecipitation followed by sequencing (ChIP-Seq) is a technique for genome-wide profiling of DNA-binding proteins, histone modifications, or nucleosomes (Myers *et al.* 2015). ChIP-Seq has become an essential tool for studying gene regulation and epigenetic mechanisms offering higher resolution, less noise, and greater coverage than its array-based predecessor ChIP-chip, especially after the decreasing cost of sequencing (Park 2009; Myers *et al.* 2013). This approach has six main steps: (1) it is initiated with cell cultures that are grown in defined conditions. When cultures reached the desired stage of development, they are treated with formaldehyde for crosslinking of proteins and DNA; (2) the chromatin is sheared by sonication into small fragments (200–600 bp); (3) an antibody specific to the protein is used to immunoprecipitate the DNA-protein complex; (4) the crosslinks are reversed by heating; (5) the released DNA are subjected to high-throughput sequencing; (6) *in silico* analysis in which resulting sequencing reads are studied for quality and cropped based



on quality of the reads (Johnson et al. 2007; Robertson et al. 2007; Park 2009). The cropped reads are then aligned to a reference genome. Afterward, areas of enrichment in the ChIP-seq data is identified and represent where the TF binds throughout the genome, usually called as peaks. CisGenome, MOSAiCs and MACS are some known algorithms that have been utilized for bacterial ChIP-seq analysis (Gentleman et al. 2004; Park 2009). After peaks are associated with genes downstream, a number of bioinformatics analyses can be carried out, including identification and analysis of motifs, differential analysis and association with expression data for deep understanding of bacterial regulon as shown in Figure 3 (Myers et al. 2015).



**Figure 3.** ChIP-seq sample preparation and analysis. Adapted from Myers et al. 2015.

As the whole-genome transcription profiling cannot reveal whether the influence of the Transcription Factors (TF) on RNA levels is direct or indirect, this requires identification of Transcription Factors binding within the appropriate promoter region. The ChIP-seq provide information about where transcription factors are bind, thus, by integrating ChIP methods and transcription profiling, it is possible to identify all direct

regulatory targets of a TF for a given condition. For example, work carried out by Stringer and collaborators (2014) on *araC* gene of *Escherichia coli*, and *Salmonella enterica* identified direct regulatory targets of AraC, including 5 novel target genes: *ytfQ*, *ydeN*, *ydeM*, *ygeA*, and *polB* (Stringer et al. 2014). Though ChIP-seq has been used only in moderation to study bacterial systems in a few bacterial species, such as, *Vibrio harveyi*, *V. cholerae*, *Rhodobacter sphaeroides*, *Mycobacterium tuberculosis*, *S. enterica*, and *Caulobacter crescentus* (Kahramanoglou et al. 2011; Myers et al. 2013, 2015; Singh et al. 2014; Haycocks et al. 2015) it is used for the identification of novel regulatory interactions, even for well-studied proteins (Grainger et al. 2007).

ChIP-seq could be an efficient tool in combination with RNA-seq to get the detail information of bacterial transcription regulation and how bacteria respond to different external conditions.

### **3.5. RNA Immunoprecipitation Sequencing (RIP-seq)**

RNA Immunoprecipitation (RIP) is the study of intracellular RNA and protein binding; it is a tool to understand the dynamic process of post-transcriptional regulatory networks. In this technique, an antibody against a protein of interest is used to recover RNA species bound to the protein and since the sequence information of RNA species bound to a specific protein is often desired, an approach combining RNA immunoprecipitation with sequencing technology (RIP-seq) was created (Chu and Corey 2012). The main challenges of RIP-seq is the cross linking step, which is relatively inefficient and only small amount of RNA is available for the library construction (Hafner et al. 2010; Chu and Corey 2012). After that, treatment with endonuclease elucidates the specific binding sites within the RNA as they will be protected from digestion, followed by the RNA protein complexes purification using electrophoresis and high throughput sequencing (Chi et al. 2009; Chu and Corey 2012), and finally the data obtained from the sequencer are analyzed using bioinformatics tool. The first work using the RIP-seq based technique was carried in *Salmonella* by Sittka and collaborators (2008) (Sittka et al. 2008). In this analysis, they used RNA-binding property of Hfq protein, as result many new sRNA were discover that were not reported in previous studies (Cho et al. 2013). Thus, RIP-Seq could be an efficient tool for the identification of bacterial non-coding RNAs.

### **3.6. LEA-seq (Low Error Amplicon Sequencing)**

The LEA-seq technique (Low Error Amplicon Sequencing) emerged in 2013 and was developed and patented by Gordon and Faith (2014) (Gordon and Faith 2014). This method was created in order to improve the quality and the depth of sequencing runs, since the mass production data produced by NGS has caused a high rate of errors in the sequencing, due to algorithms or platform reading length problems (Gordon and Faith 2014).

LEA-seq is a nucleic acids sequencing technique that identifies events that occur at low frequency, seeking to understand mutation events. The three basic steps for implementing this technique are: (1) linear PCR, (2) exponential PCR and (3)

sequencing. Whereas this technique is performed based on the bacterial 16S sequencing, in which PCR carries numerous times, and each PCR amplified uses specific primers for each linear molecule (Gordon and Faith 2014).

The LEA-seq technique is a quantitative method and presents the advantages of the generation and reading, which permits the formation of a consensus and eliminating errors for each molecule. Currently, the available techniques do not enable error detection in sequencing or whether there is a really variation in the sequence of that microorganism. The multiple sequencing through the LEA-seq technique allows for better quality and precision about the organism.

The study of Faith and collaborates (2013) aimed to identify the composition of the fecal microbiota of adults, the understanding of the role of these bacterial species and theirs therapeutic potential for intestinal diseases. This technique allowed not only work with a large number of samples (over 500 isolates), as well as contribute to a fast and accurate analysis of the data (Faith et al. 2013).

Researchers have a continuous interest in improving this technique, since it can be used for clinical investigation due to their high accuracy in cases, for example, patients with genetic mutations or somatic mutations again. LEA-seq can assist in the search and knowledge of the intestinal microbiota, as it may reveal its composition, opening up prospects for diagnosis, treatment and prevention of gastrointestinal tract diseases.

### **3.7. CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)**

Ishino and collaborates (1987) described CRISPR for the first time (Ishino et al. 1987). This system has been identified in 40% of bacterial genomes so far (Kunin, Sorek and Hugenholtz 2007) and they are defined as short repetitions of grouped bases. Although, the determination of the CRISPR locus and the characterization of adjacent genes, known as *cas* genes, responsible for the function of CRISPR, only happened in 2002 (Jansen et al. 2002). The system CRISPR/Cas uses small non-coding RNAs in association with Cas proteins. The Cas9 is a nuclease, which cleaves DNA in the selected region, so that the CRISPR system/Cas9 can be used for editing genomes.

The activity of the CRISPR/Cas involves three main mechanisms: (1) acquisition, the step which the DNA fragment was inserted into the CRISPR locus in the genome of interest of the study; (2) transcription, which the CRISPR locus will be transcribed and processed; (3) interference, which occur the ejection of nucleic acids; all those mechanisms contributes for bacterial persistence in the environment (Gasiunas et al. 2012; Jinek et al. 2012). Furthermore, the CRISPR provides mechanisms to limit the spread of antibiotic resistance or virulence factors, but the study carried by Gophna and collaborates (2015) demonstrates that even through there is different measurements to evaluate the horizontal gene transfer is not possible to identify a correlation between the CRISPR/Cas system and the evolution of the species, so that changes occur only at the population level (Gophna and Ron 2003).

RNA-seq helped annotation transcription of regions, mainly non-coding, as well as enables the identification of CRISPR elements in prokaryotes (Heidrich et al. 2015).

The CRISPR system can also be used as a tool in studies centred in gene regulation, since this system is able to activate or repress genes.

The working of Zoepfel and Randau (2013) discusses how the structure of CRISPR can affect the maturation of the RNA and thus influence the functionality of the CRISPR/Cas system (Zoepfel and Randau 2013). The RNA-seq approach was used to evaluate differential gene expression in *S. aureus*, a pathogen of major importance, and was capable to identify the CRISPR in these strains, and investigating their possible role, since these regions shows an adaptive response to infection (Osmundson, Dewell and Darst 2013). Thus, we have seen how important the use of RNA-seq approach to the magnification of knowledge and function in prokaryotes.

#### 4. RNA sequencing platforms

The RNA-seq approach can be applied in different next-generation sequencing platforms and the results obtained by then are proportional to the machine capability. In Table 3 a comparison are being made with some of the most currently employed platforms (Li et al. 2014).

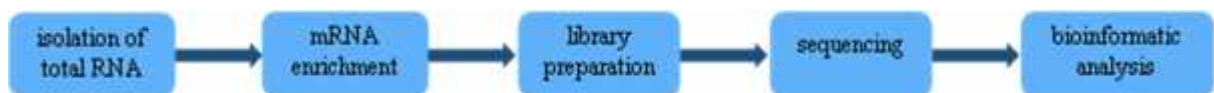
**Table 3.** Different Next Generations sequencing platforms in the study of RNA-seq. Adopted and modified from Li et al. 2014.

Company Name	Instrument	Version	Run Time (Hours)	Read Lengths (Mean)	Reads Per Run (Millions)	Applications
Illumina	HiSeq 2000	High Output	132	50	6,000	Gene expression, Splice junction detection, variant calling, fusion
Illumina	HiSeq 2500	High Output	132	50	6,000	Gene expression, Splice junction detection, variant calling, fusion
Illumina	MiSeq	v2 kit	39	250	30	Splice junction detection, variant calling,
Life Technologies	PGM	318 Chip	7.3	176	6	Splice junction detection, variant calling
Life Technologies	Proton	Proton I chip	2-4	81	70	Gene expression, Splice junction detection, variant calling
Pacific Biosciences	RS	RS	0.5-2	1,289	0.03	Splice junction detection, variant calling, full-length gene coverage
Roche	454	GS FLX+	20	686	1	Splice junction detection

## 5. Bioinformatic analysis

Experimental investigations in prokaryotes has been facilitated, extended and complemented by computational approaches (Backofen et al. 2014). Huge amounts of data has been generated by RNA-seq experiments and needs to be stored and analysed using computational techniques and tools (McGettigan 2013). This amount has become a bottleneck to bioinformatic analysis and to biologists, since today's transcriptome analysis consist of experiments and data evaluation (Backofen et al. 2014). Extract biological information from RNA-seq datasets requires bioinformatic knowledge and tools, making the software choice an important issue to successful RNA-seq analysis (Pinto et al. 2011; Backofen et al. 2014).

According to Chierico and collaborates (2015) (Del Chierico et al. 2015) and Pinto and collaborates (2011) (Pinto et al. 2011) RNA-seq process can be understood as a five-step process: (1) isolation of total RNA of the organism; (2) mRNA enrichment; (3) synthesis of cDNA; (4) NGS sequencing, which returns raw data; (5) bioinformatic analysis (Pinto et al. 2011). A flowchart of this process can be seen in Figure 4.



**Figure 4.** RNA-seq five-step process.

This session focuses on bioinformatic analysis and available computational tools. Based on literature review (Pinto et al. 2011; McClure et al. 2013b; Backofen et al. 2014; Del Chierico et al. 2015; Tjaden 2015), bioinformatic analysis can be comprehended as the extraction of biological information from the sequencing of raw data and classified/divided (Figure 5).



**Figure 5.** Bioinformatic analysis workflow.

### 5.1. Bioinformatic workflow

Quality check step aims to increase the accuracy of the results by removing sequences that can contain errors (de Sá et al. 2015), trimming sequences introduced in the library preparation step, as adapters and poly(A)-tails (Forstner, Vogel and Sharma 2014), and removing reads with low phred quality. However, regarding that, the usage of poor-quality bases can lead to less precise results (Ramos et al. 2011), and considering this, quality check can affect drastically the next steps.

Some RNA-Seq pipelines implements quality checking, like ReaDemption (Forstner, Vogel and Sharma 2014), which performs quality trimming, removes adapters

and poly(A) tails, and discards reads shorter than a given cut-off (the default cut-off is 12 nucleotides (nt)). Quality Assessment software (Ramos et al. 2011) evaluates quality based on quality-graph analysis and estimated coverage. According to Backofen and collaborators (2014) (Backofen et al. 2014), FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) is a commonly used tool to check reads quality determining the quality profile of the reads. Software suites can be also used to this purpose, FASTX-Toolkit ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)) provides tools to remove sequences attached in previous steps and other pre-processing strategies of raw data.

After quality check, if there is a reference genome mapping step will be done, otherwise de novo assembly. Mapping consists of producing the transcriptome map through aligning reads to a reference genome (Pinto et al. 2011) aiming to detect the right position of reads and distinguish between sequencing errors and genetic variations (Caboche et al. 2014). Lots of mapping softwares have been released differing in their algorithms, memory management, velocity and computational cost (Backofen et al. 2014) making the choice of a mapping tool a challenge. McClure and collaborators (2013) (McClure et al. 2013b) made a comparison between SOAP2, BWA, Bowtie and Bowtie2 aligners using 75 RNA-seq experiment data, and the comparison of mapping algorithms applied to IonTorrent data can be seen in (Caboche et al. 2014). After mapping quality can be evaluated, ReadXplorer software offers quality classification of read mapping to provide information about quality and quantity to each single read mapping (Hilker et al. 2014). This approach is recommended when a high-quality genome is available as reference, and if it is unavailable, transcripts should be assembled de novo (Tjaden 2015).

De novo assembly can be used when investigating poorly studied organisms (Haas et al. 2012), complex microbial communities or uncultivable organisms (Tjaden 2015). Both DNA and RNA must be assembled, but transcriptome assembly is significantly different than genome assembly (Haas et al. 2013), thus it is important to use RNA assemblers. Tjaden (2015) (Tjaden 2015) affirms that assemblers should be specifically designed to prokaryotes, owing to the different challenges eukaryotic and prokaryotic transcriptome faces. Bacterial genomes are often denser than eukaryotic, considering the proximity of the genes, neighbouring bacterial transcripts can overlap making it difficult to identify appropriately transcript boundaries and non-coding eukaryotic RNAs models are not appropriate to detect bacterial small regulatory RNAs (Tjaden 2015). An assembly comparison of three different softwares (Trinity, SOAPdenovo2 and Rockhooper 2) using data from 9 different bacteria can be seen in (Tjaden 2015).

When reference mapping or de novo assembly is done, data can be analysed structurally and differentially. Differential analyses main purpose is to determine the difference expression between different growth conditions or treatments (Seyednasrollah, Laiho and Elo 2013). Several softwares have been released for this purpose, but there is not a consensus about best practices, what makes difficult to choose a tool or method to use. Aiming to help in this issue Seyednasrollah and collaborators (2013) (Seyednasrollah, Laiho and Elo 2013) made a comparison among

eight differential expression software packages using two real datasets publicly available. Softwares that analyses differential expression can be based on Poisson method (DEGseq and Myrna), negative binomial method (edgeR and DEseq) or in other methods (Pinto et al. 2011; Seyednasrollah, Laiho and Elo 2013). Pinto and collaborators (2011) (Pinto et al. 2011) recommends using DEseq or edgeR when analysing replicates.

Transcriptome annotation and classification can be accentuated in structural analysis, evaluating transcripts regarding the genomic region they have been associated to and classified in: protein-coding, non-coding and intergenic regions (Backofen et al. 2014). Aiming to predict ncRNA transcripts several computational methods has been developed, Herbig and Nieselt (2011) (Herbig and Nieselt 2011) highlights SIPHT, sRNAFinder, sRNAscanner, NOCORNAr and sRNAPredict softwares. NOCORNAr distinguish itself by prediction and characterization of ncRNAs in bacteria (Herbig and Nieselt 2011).

Assess transcripts concerning genomic regions relies on transcript annotation. Computational approach is convenient to be used due its velocity and precision compared to manual annotation; however, human supervision of the results is considered important to avoid false-positives or missing features (Creecy and Conway 2015). In this technique some main structures must be detected: 5' transcript ends, 3' transcript ends, TSS and operon (Backofen et al. 2014; Creecy and Conway 2015).

#### **a. Transcript boundaries identification**

Annotation of transcript boundaries is important to operon identification and regulatory analyses (Creecy and Conway 2015). Identifying 5' UTR is not always possible, a significant number of transcripts lacking 5' UTR were found in bacteria and called leaderless transcripts, in this, transcripts translation start site and transcription start site remains in almost the same position (Backofen et al. 2014). Annotation of 3' UTR is important to get the whole analytical value of RNA-seq data, Creecy and Conway (2014) (Creecy and Conway 2015) affirm that current best 3' ends detection method is to search correlation between replicates data, highlighting that, the software package TransTermHP can find intrinsic terminators successfully.

#### **b. TSS identification**

TSS annotation can assist ncRNA annotation and polycistronic transcripts (Backofen et al. 2014); according to Creecy and Conway (2013) (Creecy and Conway 2015), it is essential to discover unknown transcripts and to analyse operon, 5' UTR and promoters architecture. Although, there is not a well-established strategies to TSS identification, owing scarce knowledge about transcription start sites in bacteria Due to computational developments in both computational analyses and “wet-lab” experiments, TSS annotation becomes feasible (Backofen et al. 2014). TSSAR is a dRNA-seq data-based tool for rapid annotation of TSS that considers dRNA-seq library statistics (Amman et al. 2014), and according to Backofen and collaborators (2014) (Backofen et al. 2014), the main advantage remains in statistical analysis presented as an easy-to-use web service. TSSpredator tool presents an automated TSS detection and

classification from RNA-seq data performing a genome wide comparative prediction of TSS (Dugar et al. 2013. A comparison among manual annotation, TSSpredator and TSSAR annotation can be seen in (Amman et al. 2014).

### c. Operon identification

The operon represents clusters of co-transcribed genes regulated by the same regulatory sequence and co-transcribed in a single mRNA. This structure has a huge biological importance improving functional gene annotation and giving important information to studies in drug targeting, functional analyses and antibiotic resistance (Chuang et al. 2012). To handle operon occurrence complexity, it should be detected by operon architecture (i.e. 5' ends and 3' ends) and have a sufficient read coverage to connect promoters and terminators. A strong indication that an operon is real is if at least 90% bases of the reads is covered (Creecy and Conway 2015). Chuang and collaborates (2012) (Chuang et al. 2012) classify computational methods to predict operons and evaluates 15 algorithms in respect to accuracy, specificity and sensitivity.

## 5.2. RNA-seq Pipeline tools

Not all pipeline tools contemplate hole RNA-seq workflow described earlier, to help in the tool choice a software functionalities comparison were developed and can be seen in Table 4. To additional support, important issues about each softwares are described forward.

**Table 4.** Software comparison.

Tool	Quality Check	Mapping	<i>De novo</i> assembly	Differential analyses
Rockhopper (McClure et al. 2013)	-	x	x	x
Rockhopper 2 (Tjaden B. 2015)	-	-	x	x
RNA-Rocket (Warren et al. 2015)	x	x	-	x
READemption (Förstner, Vogel, Sharma 2014)	x	x	-	x
ReadXplorer (Hilker et al. 2014)	x	-	-	x

Rockhopper is a system designed specifically to bacterial transcriptome RNA-seq data analyses. A novel approach to map transcripts is implemented in this software (similar to Bowtie2 approach). Mapping normalization is performed followed by transcripts assembly, identification of transcript boundaries, quantification of transcript abundance, testing for differential gene expression and operon prediction.



Analysis results are presented using Integrative Genome Viewer, allowing different experiments to be viewed simultaneously (McClure et al. 2013b).

Rockhopper 2 is a comprehensive system focused in *de novo* assembly that supports differential analyses and transcripts abundance quantification, according to Tjaden (2015) (Tjaden 2015) it does not require high-performance computers and can run on personal computers. Rockhopper 2 implements a novel *de novo* assembly algorithm to bacterial transcriptomes, the algorithm works in two stages: (1) candidate transcripts are assembled using a found k-mer; (2) sequencing reads are mapped to candidate transcripts aiming to filter candidate transcripts to high-quality final transcripts. Concerning differential analyses, Rockhopper 2 first normalize each RNA-seq dataset enabling to compare different experiments or samples (Tjaden 2015).

RNA-Rocket aims to simplify the process of aligning RNA-seq data to a reference genome and generate quantitative transcript profiles. It is built on Galaxy to provide the necessary tools and services to process RNA-seq data, some benefits of its usage are the possibility of sharing results across research groups, the support of batch analyses for multiple samples and the integration of tools and projects integrating data from the PATRIC platform (Warren et al. 2015).

READemption pipeline aims to integrate individual RNA-seq analyses tasks and putting it into a friendly tool with a command line interface, this tool was primarily developed to analyse bacterial transcriptome. In order to use full modern computers capacity and reduces run time READemption offers parallel data processing. First performs quality polyA and adapters trimming followed by mapping, coverage calculation, gene expression quantification, differential gene expression analyses and plotting. The software is able to analyse RNA-seq data from Illumina and 454 platforms.

ReadXplorer offers straightforward visualizations and analysis functions built around its unique read mapping classification. Analysis as TSS and operon detection, differential expression, RPKM value and read count calculations are available in ReadXplorer and can be exported to Microsoft Excel files. Read mapping classification sorts read mappings into three different classes: perfect match, best match and common match. This classification is incorporated in all analyses functions.

### **5.3. Bioinformatic Challenges**

Through bibliographic research (Dai et al. 2012; Mc Clure et al. 2013b; McGettigan 2013; Van Verk et al. 2013; Forstner, Vogel and Sharma 2014; Tjaden 2015), it has been concluded that bioinformatics has many challenges related to computational issues. RNA-seq experiments generate large amounts of data that must be computationally processed, analyzed, stored and retrieved using a great deal of computational power. In addition to the computational issues, it is important to take into account that not all bioinformatic researchers have extensive computational experience: this makes the lack of user-friendly tools a problem for some users and an import issue for developers. However, great computers, excellent bioinformatic researchers and user friendly tools do not guarantee successful analysis. The software selected must be appropriate to each biological question and to the organisms studied. Even with all

questions presented here, RNA-seq analysis has been very successful in recent years. This success can lead us to imagine the wonderful possibilities for RNA-seq bioinformatic analyses in the future.

## References

1. Creecy JP, Conway T. Quantitative bacterial transcriptomics with RNA-seq. *Curr Opin Microbiol* 2015;23:133–40.
2. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 2005 Sep 15;437(7057):376–80.
3. Perkins TT, Kingsley RA, Fookes MC, Gardner PP, James KD, Yu L, et al. A strand specific RNA-Seq analysis of the transcriptome of the typhoid bacillus *Salmonella typhi*. *PLoS Genet* 2009 Jul;5(7):e1000569.
4. Passalacqua KD, Varadarajan A, Ondov BD, Okou DT, Zwick ME, Bergman NH. Structure and complexity of a bacterial transcriptome. *J Bacteriol* 2009 May 15;191(10):3203–11.
5. Pinto AC, Sá PHCG de, Ramos RTJ, Barbosa S, Barbosa HPM, Ribeiro AC, et al. Differential transcriptional profile of *Corynebacterium pseudotuberculosis* in response to abiotic stresses. *BMC Genomics* 2014 Jan 9;15(1):14.
6. Westermann AJ, Gorski, SA and Vogel J. Dual RNA-seq of pathogen and host. *Nat Rev Microbiol* 2012;10:618–30.
7. Macklaim JM, Fernandes AD, Bella JMD, Hammond J-A, Reid G, Gloor GB. Comparative meta-RNA-seq of the vaginal microbiota and differential expression by *Lactobacillus iners* in health and dysbiosis. *Microbiome* 2013 Apr 12;1(1):12.
8. Leimena MM, Ramiro-Garcia J, Davids M, van den Bogert B, Smidt H, Smid EJ, et al. A comprehensive metatranscriptome analysis pipeline and its validation using human small intestine microbiota datasets. *BMC Genomics* 2013;14:530.
9. Bisanz JE, Macklaim JM, Gloor GB, Reid G. Bacterial metatranscriptome analysis of a probiotic yogurt using an RNA-Seq approach. *Int Dairy J* 2014;39(2):284–92.
10. Wiegand S, Dietrich S, Hertel R, Bongaerts J, Evers S, Volland S, et al. RNA-Seq of *Bacillus licheniformis*: active regulatory RNA features expressed within a productive fermentation. *BMC Genomics* 2013;14(1):667.
11. Wang Z, Yang S-T. Propionic acid production in glycerol/glucose co-fermentation by *Propionibacterium freudenreichii* subsp. *shermanii*. *Bioresour Technol* 2013;137:116–23.
12. Sorek R, Cossart P. Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nat Rev Genet* 2010 Jan;11(1):9–16.
13. Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiß S, Sittka A, et al. The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature*. 2010 Mar 11;464(7286):250–5.
14. Haas BJ, Chin M, Nusbaum C, Birren BW, Livny J. How deep is deep enough for RNA-Seq profiling of bacterial transcriptomes? *BMC Genomics* 2012 Dec 27;13(1):734.

15. Bischler T, Siew Tan H, Nieselt K, Sharma CM. Differential RNA-seq (dRNA-seq) for annotation of transcriptional start sites and small RNAs in *Helicobacter pylori*. *Methods* [Internet]. 2015 Jul 6 [cited 2015 Jul 6]; Available from: <http://www.sciencedirect.com/science/article/pii/S1046202315002546>
16. Kratz A, Carninci P. The devil in the details of RNA-seq. *Nat Biotechnol* 2014 Sep;32(9):882–4.
17. Sandler E, Johnson GD, Krawetz SA. Local and global factors affecting RNA sequencing analysis. *Anal Biochem* 2011;419(2):317–22.
18. Liu Y, Zhou J, White KP. RNA-seq differential expression studies: more sequence or more replication? *Bioinformatics* 2014 Feb 1;30(3):301–4.
19. Isabella VM, Clark VL. Deep sequencing-based analysis of the anaerobic stimulon in *Neisseria gonorrhoeae*. *BMC Genomics* 2011 Jan 20;12(1):51.
20. Patenge N, Pappesch R, Khani A, Kreikemeyer B. Genome-wide analyses of small non-coding RNAs in streptococci. *Front Genet* 2015;6:189.
21. Gottesman S, Storz G. Bacterial small RNA regulators: versatile roles and rapidly evolving variations. *Cold Spring Harb Perspect Biol* [Internet]. 2011 Dec [cited 2015 Jul 2];3(12). Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3225950/>
22. Papenfort K, Vogel J. Regulatory RNA in bacterial pathogens. *Cell Host Microbe* 2010 Jul 22;8(1):116–27.
23. Kröger C, Dillon SC, Cameron ADS, Papenfort K, Sivasankaran SK, Hokamp K, et al. The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Typhimurium. *Proc Natl Acad Sci* 2012 May 15;109(20):E1277–86.
24. Yan Y, Su S, Meng X, Ji X, Qu Y, Liu Z, et al. Determination of sRNA expressions by RNA-seq in *Yersinia pestis* grown in vitro and during infection. *PLoS ONE* 2013 Sep 11;8(9):e74495.
25. Storz G, Vogel J, Wassarman KM. Regulation by small RNAs in bacteria: expanding frontiers. *Mol Cell* 2011 Sep 16;43(6):880–91.
26. Van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C. Ten years of next-generation sequencing technology. *Trends Genet* 2014;30(9):418–26.
27. Humphrys MS, Creasy T, Sun Y, Shetty AC, Chibucos MC, Drabek EF, et al. Simultaneous transcriptional profiling of bacteria and their host cells. *PLoS ONE* 2013 Dec 4;8(12):e80597.
28. Martin JA, Wang Z. Next-generation transcriptome assembly. *Nat Rev Genet* 2011 Sep 7;12(10):671–82.
29. Tjaden B. De novo assembly of bacterial transcriptomes from RNA-seq data. *Genome Biol* 2015;16(1):1.
30. Pinto AC, Ramos RTJ, Silva WM, Rocha FS, Barbosa S, Miyoshi A, et al. The core stimulon of *Corynebacterium pseudotuberculosis* strain 1002 identified using ab initio methodologies. *Integr Biol* 2012;4(7):789.
31. Innocenti N, Golumbeanu M, d' Hérœuël AF, Lacoux C, Bonnin RA, Kennedy SP, et al. Whole genome mapping of 5'RNA ends in bacteria by tagged sequencing: a comprehensive view in *Enterococcus faecalis*. *ArXiv Prepr ArXiv14101925* [Internet]. 2014 [cited 2014 Dec 15]; Available from: <http://arxiv.org/abs/1410.1925>

32. Fouquier d'Herouel A, Wessner F, Halpern D, Ly-Vu J, Kennedy SP, Serror P, et al. A simple and efficient method to search for selected primary transcripts: non-coding and antisense RNAs in the human pathogen *Enterococcus faecalis*. *Nucleic Acids Res* 2011 Apr 1;39(7):e46–e46.
33. Mamanova L, Turner DJ. Low-bias, strand-specific transcriptome Illumina sequencing by on-flowcell reverse transcription (FRT-seq). *Nat Protoc* 2011 Nov;6(11):1736–47.
34. Mamanova L, Andrews RM, James KD, Sheridan EM, Ellis PD, Langford CF, et al. FRT-seq: amplification-free, strand-specific transcriptome sequencing. *Nat Methods* 2010 Feb;7(2):130–2.
35. Vergara-Irigaray M, Fookes MC, Thomson NR, Tang CM. RNA-seq analysis of the influence of anaerobiosis and FNR on *Shigella flexneri*. *BMC Genomics* 2014 Jun 6;15(1):438.
36. Myers KS, Park DM, Beauchene NA, Kiley PJ. Defining bacterial regulons using ChIP-seq methods. *Methods* [Internet]. 2015 [cited 2015 Jul 17]; Available from: <http://www.sciencedirect.com/science/article/pii/S1046202315002285>
37. Myers KS, Yan H, Ong IM, Chung D, Liang K, Tran F, et al. Genome-scale analysis of *Escherichia coli* FNR reveals complex features of transcription factor binding. *PLoS Genet* 2013;9(6):e1003565.
38. Park PJ. ChIP-seq: advantages and challenges of a maturing technology. *Nat Ver Genet* 2009;10(10):669–80.
39. Johnson DS, Mortazavi A, Myers RM, Wold B. Genome-wide mapping of in vivo protein-DNA interactions. *Science* 2007;316(5830):1497–502.
40. Robertson G, Hirst M, Bainbridge M, Bilenky M, Zhao Y, Zeng T, et al. Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat Methods* 2007;4(8):651–7.
41. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 2004;5(10):R80.
42. Stringer AM, Currenti S, Bonocora RP, Baranowski C, Petrone BL, Palumbo MJ, et al. Genome-scale analyses of *Escherichia coli* and *Salmonella enterica* AraC reveal noncanonical targets and an expanded core regulon. *J Bacteriol* 2014;196(3):660–71.
43. Haycocks JRJ, Sharma P, Stringer AM, Wade JT, Grainger DC. The molecular basis for control of ETEC enterotoxin expression in response to environment and host. *PLoS Pathog* 2015 Jan 8;11(1):e1004605.
44. Singh SS, Singh N, Bonocora RP, Fitzgerald DM, Wade JT, Grainger DC. Widespread suppression of intragenic transcription initiation by H-NS. *Genes Dev* 2014 Feb 1;28(3):214–9.
45. Kahramanoglou C, Seshasayee ASN, Prieto AI, Ibberson D, Schmidt S, Zimmermann J, et al. Direct and indirect effects of H-NS and Fis on global gene expression control in *Escherichia coli*. *Nucleic Acids Res* 2011 Mar 1;39(6):2073–91.

46. Wade JT, Struhl K, Busby SJ, Grainger DC. Genomic analysis of protein–DNA interactions in bacteria: insights into transcription and chromosome organization. *Mol Microbiol* 2007;65(1):21–6.
47. Grainger DC, Aiba H, Hurd D, Browning DF, Busby SJW. Transcription factor distribution in *Escherichia coli*: studies with FNR protein. *Nucleic Acids Res* 2007 Jan 1;35(1):269–78.
48. Chu Y, Corey DR. RNA sequencing: platform selection, experimental design, and data interpretation. *Nucleic Acid Ther* 2012;22(4):271–4.
49. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, et al. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 2010;141(1):129–41.
50. Chi SW, Zang JB, Mele A, Darnell RB. Argonaute HITS-CLIP decodes microRNA–mRNA interaction maps. *Nature* 2009;460(7254):479–86.
51. Sittka A, Lucchini S, Papenfort K, Sharma CM, Rolle K, Binnewies TT, et al. Deep sequencing analysis of small noncoding RNA and mRNA targets of the global posttranscriptional regulator, Hfq. *PLoS Genet* 2008;4(8):e1000163.
52. Cho S, Cho Y, Lee S, Kim J, Yum H, Kim SC, et al. Current challenges in bacterial transcriptomics. *Genomics Inform* 2013;11(2):76.
53. Gordon JI, Faith JJ. Methods of low error amplicon sequencing (LEA-Seq) and the use thereof [Internet]. Google Patents; 2014 [cited 2015 Jul 14]. Available from: <https://www.google.com/patents/US20140357499>
54. Faith JJ, Guruge JL, Charbonneau M, Subramanian S, Seedorf H, Goodman AL, et al. The long-term stability of the human gut microbiota. *Science* 2013;341(6141):1237439.
55. Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol* 1987;169(12):5429–33.
56. Kunin V, Sorek R, Hugenholtz P. Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol* 2007;8(4):R61.
57. Jansen R, Embden J, Gaastra W, Schouls L, others. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* 2002;43(6):1565–75.
58. Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci* 2012;109(39):E2579–86.
59. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *Science* 2012;337(6096):816–21.
60. Gophna U, Ron EZ. Virulence and the heat shock response. *Int J Med MicrobiolIJMM* 2003 Feb;292(7-8):453–61.
61. Heidrich N, Dugar G, Vogel J, Sharma CM. Investigating CRISPR RNA biogenesis and function using RNA-seq. *CRISPR Methods Protoc* 2015;1–21.
62. Zoephel J, Randau L. RNA-Seq analyses reveal CRISPR RNA processing and regulation patterns. *Biochem Soc Trans* 2013 Dec;41(6):1459–63.

63. Osmundson J, Dewell S, Darst SA. RNA-Seq reveals differential gene expression in *Staphylococcus aureus* with single-nucleotide resolution. *PLoS ONE* 2013 Oct 7;8(10):e76572.
64. Li S, Tighe SW, Nicolet CM, Grove D, Levy S, Farmerie W, et al. Multi-platform assessment of transcriptome profiling using RNA-seq in the ABRF next-generation sequencing study. *Nat Biotechnol* 2014 Aug 24;32(9):915–25.
65. Backofen R, Amman F, Costa F, Findei S, Richter AS, Stadler PF. Bioinformatics of prokaryotic RNAs. *RNA Biol* 2014;11(5):470–83.
66. McGettigan PA. Transcriptomics in the RNA-seq era. *Curr Opin Chem Biol* 2013 Feb; 17(1):4–11.
67. Pinto AC, Melo-Barbosa HP, Miyoshi A, Silva A, Azevedo V. Review application of RNA-seq to reveal the transcript profile in bacteria. *Genet Mol Res* 2011;10(3):1707–18.
68. Del Chierico F, Ancora M, Marcacci M, Camma C, Putignani L, Conti S. Choice of next-generation sequencing pipelines. *Bacterial Pangenomics* [Internet]. Springer; 2015 [cited 2015 Jul 14]. p. 31–47. Available from: [http://link.springer.com/protocol/10.1007/978-1-4939-1720-4\\_3](http://link.springer.com/protocol/10.1007/978-1-4939-1720-4_3)
69. McClure R, Balasubramanian D, Sun Y, Bobrovskyy M, Sumbly P, Genco CA, et al. Computational analysis of bacterial RNA-Seq data. *Nucleic Acids Res* 2013 Aug 1;41(14):e140–e140.
70. De Sá PH, Veras AA, Carneiro AR, Pinheiro KC, Pinto AC, Soares SC, et al. The impact of quality filter for RNA-Seq. *Gene* 2015;563(2):165–71.
71. Förstner KU, Vogel J, Sharma CM. READemption—A tool for the computational analysis of deep-sequencing-based transcriptome data. *Bioinformatics* 2014;btu533.
72. Ramos RT, Carneiro AR, Baumbach J, Azevedo V, Schneider MP, Silva A. Analysis of quality raw data of second generation sequencers with Quality Assessment Software. *BMC Res Notes* 2011 Apr 18;4(1):130.
73. Caboche S, Audebert C, Lemoine Y, Hot D. Comparison of mapping algorithms used in high-throughput sequencing: application to Ion Torrent data. *BMC Genomics* 2014;15(1):264.
74. Hilker R, Stadermann KB, Doppmeier D, Kalinowski J, Stoye J, Straube J, et al. Read-Xplorer—visualization and analysis of mapped sequences. *Bioinformatics* 2014;btu205.
75. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc* 2013;8(8):1494–512.
76. Seyednasrollah F, Laiho A, Elo LL. Comparison of software packages for detecting differential expression in RNA-seq studies. *Brief Bioinform* [Internet]. 2013 Dec 2 [cited 2014 Apr 30]; Available from: <http://bib.oxfordjournals.org/cgi/doi/10.1093/bib/bbt086>
77. Herbig A, Nieselt K. nocoRNAc: characterization of non-coding RNAs in prokaryotes. *BMC Bioinformatics* 2011;12(1):40.
78. Amman F, Wolfinger MT, Lorenz R, Hofacker IL, Stadler PF, Findei S. TSSAR: TSS annotation regime for dRNA-seq data. *BMC Bioinformatics* 2014;15(1):89.

79. Dugar G, Herbig A, Förstner KU, Heidrich N, Reinhardt R, Nieselt K, et al. High-resolution transcriptome maps reveal strain-specific regulatory features of multiple *Campylobacter jejuni* isolates. 2013 [cited 2015 Jul 14]; Available from: <http://dx.plos.org/10.1371/journal.pgen.1003495>
80. Chuang L-Y, Chang H-W, Tsai J-H, Yang C-H. Features for computational operon prediction in prokaryotes. *Brief Funct Genomics* 2012;els024.
81. Warren AS, Aurrecochea C, Brunk B, Desai P, Emrich S, Giraldo-Calderón GI, et al. RNA-Rocket: an RNA-Seq analysis resource for infectious disease research. *Bioinformatics* 2015;btv002.
82. Van Verk MC, Hickman R, Pieterse CM, Van Wees SC. RNA-Seq: revelation of the messengers. *Trends Plant Sci* 2013;18(4):175–9.
83. Dai L, Gao X, Guo Y, Xiao J, Zhang Z, others. Bioinformatics clouds for big data manipulation. *Biol Direct* 2012;7(1):43.

### **Anexo 3: Artigo: Value of a newly sequenced bacterial genome**

### **Annexe 3 : Article : Value of a newly sequenced bacterial genome**

WJBC - World Journal of Biological Chemistry

Submit a Manuscript: <http://www.wjgnet.com/esps/> World J Biol Chem 2014 May 26; 5(2): 161-168 Help Desk: <http://www.wjgnet.com/esps/helpdesk.aspx> ISSN 1949-8454 (online)

DOI: 10.4331/wjbc.v5.i2.161

© 2014 Baishideng Publishing Group Inc. All rights reserved.

### **REVIEW**

#### **Value of a newly sequenced bacterial genome**

Eudes GV Barbosa, Flavia F Aburjaile, Rommel TJ Ramos, Adriana R Carneiro, Yves Le Loir, Jan Baumbach, Anderson Miyoshi, Artur Silva, Vasco Azevedo

Eudes GV Barbosa, Flavia F Aburjaile, Anderson Miyoshi, Vasco Azevedo, Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901 MG, Brazil

Eudes GV Barbosa, Jan Baumbach, Department of Mathematics and Computer Science, University of Southern Denmark, 5230 Odense, Denmark

Flavia F Aburjaile, Yves Le Loir, INRA, UMR1253, Science et Technologie du Lait et de l'OEuf, F-35042 - Rennes, France

Rommel TJ Ramos, Adriana R Carneiro, Artur Silva, Laboratório de Polimorfismo de DNA, Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém 66075-110, Brazil

Author contributions: All authors contributed extensively to the work presented in this review.



Supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) in Brazil, processes BEX 12954-12-8 and 11517-12-3, to Barbosa EGV and Aburjaile FF

Correspondence to: Vasco Azevedo, MD, PhD, Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Av. Antônio Carlos 6627 Pampulha, Belo Horizonte 31270-901, Brazil.

vasco@icb.ufmg.br

Telephone: +55-31-34092873 Fax: +55-31-34092610

Received: December 11, 2013 Revised: January 14, 2014

Accepted: April 3, 2014

Published online: May 26, 2014

### **Abstract**

Next-generation sequencing (NGS) technologies have made high-throughput sequencing available to medium and small-size laboratories, culminating in a tidal wave of genomic information. The quantity of sequenced bacterial genomes has not only brought excitement to the field of genomics but also heightened expectations that NGS would boost antibacterial discovery and vaccine development. Although many possible drug and vaccine targets have been discovered, the success rate of genome-based analysis has remained below expectations. Furthermore, NGS has had consequences for genome quality, resulting in an exponential increase in draft (partial data) genome deposits in public databases. If no further interests are expressed for a particular bacterial genome, it is more likely that the sequencing of its genome will be limited to a draft stage, and the painstaking tasks of completing the sequencing of its genome and annotation will not be undertaken. It is important to know what is lost when we settle for a draft genome and to determine the “scientific value” of a newly sequenced genome. This review addresses the expected impact of newly sequenced genomes on antibacterial discovery and vaccinology. Also, it discusses the factors that could be leading to the increase in the number of draft deposits and the consequent loss of relevant biological information.

**Keywords:** Next-generation sequencing; Drafts; Prokaryotic genomes; Computational tools; Omics.

**Core tip:** Next-generation sequencing (NGS) technologies have made high-throughput sequencing available to medium- and small-size laboratories, culminating in a tidal wave of genomic information. The quantity of bacterial genomes has not only brought excitement to the field of genomics, it has also heightened expectations that NGS would boost antibacterial discovery and vaccine development. Although many possible drug and vaccine targets have been discovered, the success rate of genome-based analysis has remained below expectations. Furthermore, NGS has consequences for genome quality, resulting in an exponential increase in draft genome deposits in public databases. This review will address the expected impact of newly sequenced genomes on antibacterial discovery and vaccinology, as well as the impact of NGS on draft bacterial genomes.

Barbosa EGV, Aburjaile FF, Ramos RTJ, Carneiro AR, Le Loir Y, Baumbach J, Miyoshi A, Silva A, Azevedo V. Value of a newly sequenced bacterial genome. *World J Biol Chem*

Barbosa EGV et al. Value of a newly sequenced bacterial genome 2014; 5(2): 161-168

Available from:

URL: <http://www.wjgnet.com/1949-8454/full/v5/i2/161.htm>

DOI: <http://dx.doi.org/10.4331/wjbc.v5.i2.161>

## INTRODUCTION

Since its release in 2005, next-generation sequencing (NGS) has been responsible for a drastic reduction in the price of genome sequencing and for a tidal wave of genetic information [1]. NGS technologies have made high-throughput sequencing available to medium- and small-size laboratories. The new possibility of generating a large number of sequenced bacterial genomes not only brought excitement to the field of genomics but also heightened expectations that the development of vaccines and the search for new antibacterial targets would be boosted. Nevertheless, these expectations were shown to be naïve. The complexity of host-bacteria interactions and the large diversity of bacterial genetic products have been shown to play greater roles in vaccine development and antibacterial discovery [2-4].

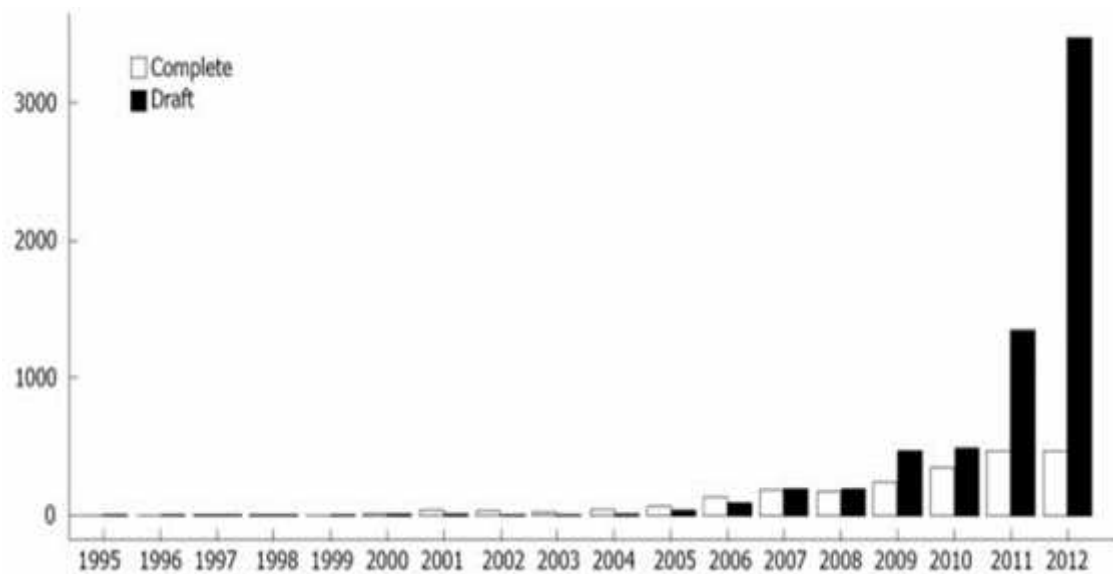
Additionally, as with any methodology, NGS presents its own drawbacks. Among the new sequencing technologies the most consolidated in the market are the 454 GS FLX platform (Roche), Illumina (Genome Analyzer) and SOLiD (Life

Technologies)[5,6]. These devices are capable of generating millions of reads, providing high coverage genomic but with a drawback, reads are considerably smaller than the ones produced by Sanger methodology[7,8]. While Sanger methodology produces reads ranging from 800 to 1000 bases, NGS platforms produces reads ranging from 50 (SOLiD V3) to  $2 \times 150$  bases (Illumina)[9]. The small amount of information contained in each read makes it difficult to completely assemble a genome using exclusively computational tools [10, 11]. Therefore small reads made the genome assembly process a quite more laborious task.

In recent years, approaches that use hybrid assemblies were developed to facilitate the assembly process. They take advantage of high read quality of second generation sequencers, i.e., Illumina (Genome Analyzer), and longer read lengths from third generation sequencers, i.e., SMRT sequencers (Pacific Biosciences) and Ion Torrent PGM[12,13]. Although empirically logical, this kind of approach wasn't facilitated due to the lack of integration between sequencers.

In order to improving and verifying quality genome is essential to know which combination of sequencing data, computer algorithms, and parameters can produce the highest quality assembly[14,15]. Also, it is necessary to know the more likely type of error data a sequencer platform will present. For instance, Illumina and SOLiD are more likely to present nucleotide substitution, while 454 GS FLX and Ion Torrent are more likely to present indels [16]. Nearly none bioinformatic system has been developed to integrate reads from different sequencers into a single assembly [12, 17]. This new developed approaches aim to reduce the manual intervention in finishing genomes, since repetitive regions may be solved using an hybrid approach.

Although NGS is directly responsible for considerable growth in the size of genomic databases, it has also been indirectly responsible for a decrease in genome quality [1, 10]. The number of draft genome (partial data) deposits in public databases has grown exponentially since 2005 (Figure 1). In general, if no further studies will be developed using a particular organism's genome, it is more likely to be deposited as a draft genome. Otherwise, the painstaking tasks of improving and finishing the genome (complete data) must be undertaken [18].



**Figure 1. Number of complete genome and draft genome (partial data) deposits in public databases.**

This review will address the “scientific value” of a newly sequenced genome and the amount of insight it can provide. We will address the factors that could be leading to the increase in the number of draft deposits and the consequent loss of relevant biological information. Additionally, we will summarize the expectations created by NGS technologies regarding vaccine development and antibacterial discovery.

### **Overview of Sequencing and assembly**

For 30 years, sequencing technologies based on Sanger chemistry dominated the market. Although sequencing had undergone numerous improvements over the years, gene cloning techniques were still necessary to obtain genomic DNA sequences. Therefore, the time and cost required to obtain a complete genome sequence remained high. Moreover, the capacity of parallel sequencing was quite limited [19-21]. NGS platforms made it possible to sequence complete prokaryotic genomes using massively parallel sequencing more rapidly and at a lower cost [20, 22].

Although NGS has facilitated sequencing processes, its relatively smaller reads make the assembly process a computational challenge [10, 11]. The main limitation of short-read assembly methods is their inability to resolve repetitive regions of the genome without paired libraries [11]. The assembly of repetitive regions was an important issue even before the introduction of NGS platforms; shorter reads only made the problem worse.

In 2001, Kececioglu et al [23] argued about the impossibility of correctly assembling regions of the genome that contain identical copies of a sequence. Usually, long DNA repeats are not exact copies. They contain small differences that could, in principle, permit their correct assembly. Nevertheless, a major difficulty arises from sequencing errors. Assembly software must accept imperfect sequencing alignments to avoid missing genuine connections between sequences [22]. With the small amount of information within each read adding to the inherent sequencing error, it is difficult to separate true differences within repeated sequences from sequencing errors.

A study by Phillippy et al [24] revealed that the majority of contig ends in draft genomes were associated with repeated regions. They concluded that it was possible to categorize the majority of mis-assembly events into two general classes: (1) repeat collapse or expansion; and (2) sequence rearrangement and inversion. Each of these classes exhibits specific mis-assembly signatures: the first class is the result of incorrect assembly in repetitive regions, including fewer or additional copies; the second class is the result of the rearrangement of multiple repeated copies, which is caused by the insertion of a read between them. The second class may be considered more influential because, if not fixed, it might be interpreted as a real biological rearrangement event [25, 26]. If the assembler cannot resolve the region between two genomic fragments, a gap is formed. Gaps may occur due to: (1) an intrinsic characteristic of the sequencing platform that leads to incomplete or incorrect information; or (2) the inability of an assembly algorithm to handle regions of low complexity or repeated DNA [18, 27, 28]. The process of identifying and closing these gaps is quite laborious and requires additional manual intervention.

Gap closure processes usually involve the design of primers flanking the gap region to perform semiautomated sequencing of the unrepresented parts of the genome [28]. Several bioinformatics methodologies have been developed to facilitate gap closure. IMAGE is a tool that uses de Bruijn methodology to fill gaps with short reads that are aligned with flanking regions of the gap and were not used in the assembly [28]. In 2011, Cerdeira et al [29] generated a similar strategy by using CLC Genomics Workbench for the recursive alignment of unused short reads from the SOLiD platform. GapFiller is another tool that uses local alignment; its main advantage is the use of paired reads to estimate gap size and allows define the type of paired library: reverse-reverse, forward-forward, reverse-forward and forward-reverse [30].

From a purely practical standpoint, assembly tools are not required to produce a perfectly finished genome as an output. Their main function is to reduce the sequencing reads to a manageable number of contigs [26]. The process of finishing a genome, ensuring that gaps are closed and the gene order is correct, requires human decisionmaking. Therefore, the lack of fully automated processes constitutes a bottleneck in generating complete genomes.

### **“Scientific Value” of a Newly sequenced genome**

The value of a newly sequenced genome can be assessed using many different metrics. If publications are considered the main “currency” within the scientific community, there has been a considerable decrease in the value of new sequences over the last four decades.

The introduction of Sanger methodology in 1977 was one of the main landmarks in the early stages of the genomic era[31]. During the first years of using Sanger sequencing, a sequence of no more than 1000 nucleotides was sufficient for a work to be accepted in a journal such as *Cell* (current impact factor: 32.40) or *Nature* (current impact factor: 36.28)[32-34]. In 1980, the shotgun DNA sequencing methodology was introduced, enabling the sequencing of longer DNA fragments [35]. Complete bacterial operons were sequenced and published in journals such as *Molecular Microbiology* (current impact factor: 5.01) and *Proceedings of the National Academy of Sciences* (PNAS -current impact factor: 9.68) [36-38].

A combination of DNA sequencing improvements and the newly developed TIGR Assembler [39] culminated in the publication of the first complete bacterial genomes in 1995. Papers containing the complete nucleotide sequences of *Haemophilus influenzae* Rd (1830137 base pairs) and *Mycoplasma genitalium* (580070 base pairs) were both published in *Science* (current impact factor: 31.20)[40,41].

Almost 20 years later, a paper containing the sequence of a prokaryotic genome alone may be published in the Genome Announcement section of the *Journal of Bacteriology* (current impact factor: 3.82) or in *Standards in Genomic Sciences* (SIGS - has not been published sufficiently long to receive an impact factor). A recent article by Smith even refers to the not-so-distant “death” of the “genome paper”, noting that the space for genome publication may come to an end soon [42].

The publication impact of newly sequenced genomes decreased following DNA sequencing improvements, and the reason is no mystery. High-impact journals only

publish groundbreaking original scientific research or results of outstanding scientific importance. To produce a higher-impact publication, more information must be extracted from genomes. For instance, several genomes may be examined in a comparative genomic analysis or pangenomic study[43,44], or an analysis may focus on the presence or absence of specific markers or on small differences between DNA sequences[26,45]. In this context, the genome becomes a stepping stone to the main goal, the comparative analysis. As the basis of the analysis, the genome sequence remains important. Nevertheless, it may not be of sufficient importance for one to undertake the painstaking task of completing the genome sequence.

### **What Is Lost When we Opt for a Draft Genome?**

Over the years, arguments have been presented in favor both of complete genomes[41,46] and of the superior “tradeoff ” that a draft genome represents[47]. The discussion has been centered around two main points: (1) to provide the greatest amount of useful data, sequences must be as complete as possible; and (2) draft genomes (partial data) are sufficient for most scientific contexts.

The issue at stake is the extra money and manpower necessary to finish a genome. Is the additional information contained in a finished genome worth the investment? To answer this question, one must identify the information that is lost from a draft and analyze the quality of data that is generated using drafts. Furthermore, it is necessary to understand the limits of draft genome use.

The first issue to consider is whether it is possible to properly identify all of an organism’s genes in a draft genome. Gene characterization consists of the following: (1) gene prediction with the identification of an open reading frame (ORF); and (2) the functional annotation of the gene product. The main gene identification problems in drafts are associated with the partial or complete loss of ORFs [10]. Such errors may lead either to over-annotation, due to the annotation of multiple fragments originating from the same ORF, or to under-annotation, possibly due to the absence of partial or entire domains from the ORF [10]. These problems affect genomic analyses, causing errors due to missing ORFs that are not annotated or due to multiple fragments that belong to the same ORF but are annotated separately. In other words, the mere absence of a gene from a draft cannot be considered definitive proof of its absence from the organism’s genome[10,41].

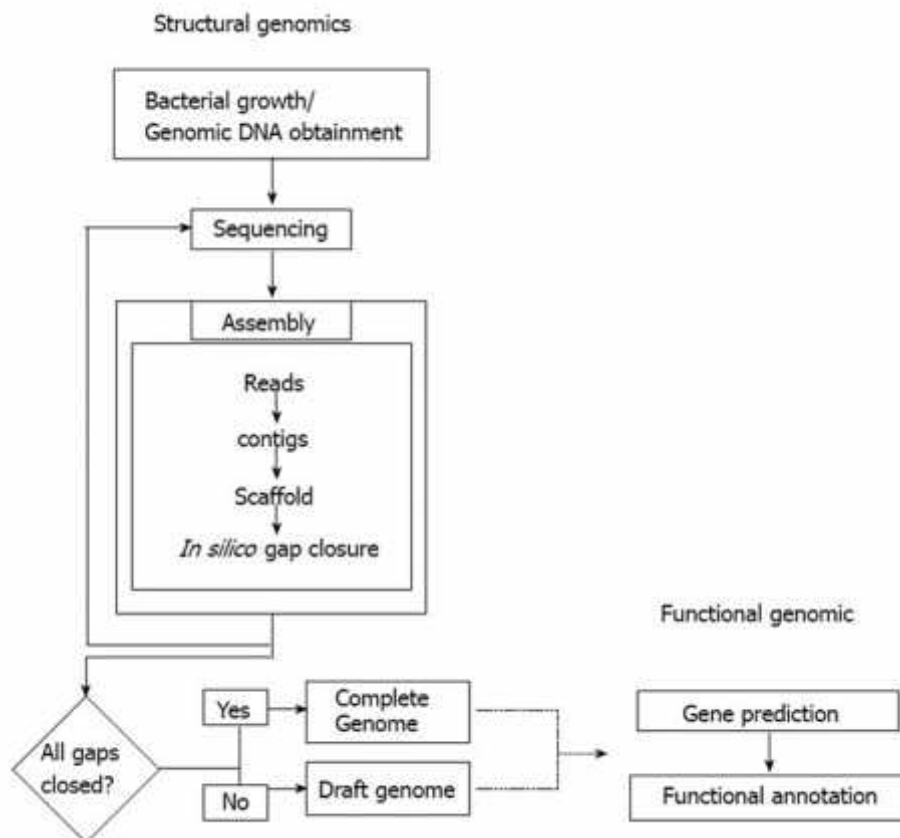
The pangenomic approach is one type of analysis that may be impaired by reliance on draft genomes, because many genes in a draft may be misidentified due to fragmentation.

Pangenomic projects attempt to characterize the gene pool of a bacterial species as the genes that are present in all strains (the “core genome”) and the genes that are present in only a few species (the “dispensable genome”)[43].

Horizontal gene transfer (HGT) analysis is another approach that cannot be performed using drafts. HGT is one of the main sources of variability among bacteria because it allows the acquisition of several new genes [36,37]. There is evidence that most gaps in genomic sequences are associated with transposases, insertion sequences and integrases, structures that usually flank a genomic island [48]. Another approach that may be impaired by reliance on drafts is phylogenomics, which aims to reconstruct both the vertical and lateral gene transfer processes of a bacterial species using a whole-genome analysis [49].

Although not strictly related to drafts, the functional annotation of genes is another feature that is usually neglected when we opt for a draft genome (Figure 2). Complete genomes may also present this problem because the quality of functional annotation is related to the amount of effort dedicated to a genome. DNA sequence is being generated much more rapidly than it can be analyzed; thus, a large proportion of the sequence information in databases has been annotated solely by automatic algorithms [50]. It is disturbing that although automatic annotation algorithms have improved over the years, misannotation has increased over time [50]. The misannotation of a reference strain is particularly harmful because the error will likely be propagated to other genomes. In our attempts to exploit the full potential of NGS, we risk having databases filled with incomplete and/or incorrect genomic data.





**Figure 2. General workflow during sequencing process a bacterial genome.**

Because the purpose of many sequencing projects is to identify a small number of differences between a newly sequenced genome and the sequence of a closely related species, a large number of genomes are left as drafts [26]. Considering the constant evolution of organisms, a sequenced genome represents a snapshot in the biological history of a species. Therefore, a single finished genome might be useful for decades of future studies. By opting for draft genomes, we may be shutting down the full gamut of future scientific analysis.

## Vaccine Development

Genomic information was expected to boost vaccine discovery. In an attempt to measure the impact of genomic information on this field, Prachi et al[2] analyzed all the patent applications that contained genomic information.

They observed that there was an enormous increase in such applications shortly after the first complete genomes were released, but since 2002, there has been a continuous decrease. The authors attributed this decrease to more stringent legal requirements, which call for empirical evidence to complement in silico data.

The initial increase in patent applications containing genomic information was related to the development of a new paradigm in vaccine development. In 2000, Rappuoli [51] described the “reverse vaccinology” (RV) concept, in which he proposed inverting the traditional process of antigen identification. Instead of identifying the antigenic components of a pathogenic organism using serological or biochemical methods, RV uses the organism’s genome to predict all of its protein antigens. RV approaches mainly focus on secreted proteins because they are more likely to induce immune responses. Secreted proteins are involved in several processes that modulate the host-pathogen relationship, such as cell adhesion and

invasion, as well as resistance to stress conditions[52-54].

Over the years, several methodologies have been developed to predict secreted proteins and to evaluate their potential immunological properties.

In 2010, Vaxign was released as the first vaccine design tool with a web interface (<http://www.violinet.org/vaxign/>). Vaxign allows users to submit their own sequences to perform vaccine target predictions. The Vaxign predictions have been consistent with existing reports for organisms such as *Mycobacterium tuberculosis* and *Neisseria meningitidis* [55]. Another vaccine design tool is MED (Mature Epitope Density -<http://med.mmci.uni-saarland.de/>). MED attempts to select the more promising vaccine targets by identifying proteins with higher concentrations of epitopes [56]. There are also tools exclusively for protein epitope prediction, such as Immune Epitope Analysis (<http://tools.immuneepitope.org/main/>) and Vaxitope (<http://www.violinet.org/vaxign/vaxitop/index.php>). Because a large number of bacterial genomes are already available, reverse vaccinology is quite accessible and inexpensive. Nevertheless, as has been previously discussed [57,58], the expectations for reverse vaccinology techniques do not correspond to reality, given the small number of vaccines have been developed using the bacterial genome sequences available[59]. This occurs because there are also several factors that are involved in the host response during infection, for example, the production of antibodies by the immune system.

### **Antibacterial Discovery**

The period between the 1930s and the 1960s is known as the “golden age” of antibiotic discovery [11, 60]. During this period, most of the known classes of antibiotics were discovered. These discoveries involved screening natural products regardless of their mechanisms of action. After most of the low-hanging fruits were

harvested, the rate of antibacterial discovery decreased, culminating in a slowdown beginning in the 1990s [61].

Hopes for turning this void into a rapid acceleration accompanied the completion of the first bacterial genome sequences. The goal was to use comparative genomic analysis to identify potential targets present in a desirable spectrum (e.g., the bacteria responsible for upper respiratory tract infections) [3, 4, 62]. It was naive to assume that having the genome sequences would be sufficient for this level of discovery; a possible drug target must undergo numerous stages, from discovery through human clinical tests, and it is not possible to develop drugs for all potential targets [3,62]. Nevertheless, the prospect of exploring hundreds of potential targets revived the interest of pharmaceutical companies.

After some years of trials, several companies ended their target-based programs because of a lack of productivity. Despite reports of multi-resistant bacterial strains, the efforts to discover new antibacterial targets were again reduced [63, 64]. Although genomics has not been able to reverse the lack of new antibiotic development, it has significantly improved screening methodologies. Genomics has facilitated high-throughput drug campaigns, which are being used to determine the mechanisms of action of antibacterial compounds and bacterial resistance mechanisms [4].

### **Conclusion**

Several next-generation platforms have been developed in recent decades, as well as bioinformatics programs to an enhancement of performance and optimization omics techniques. Is not yet possible to integrate reads from different sequencers into a single assembly [17,23]. This newly developed approach aims to reduce the amount of manual intervention needed to complete a genome sequence by using a hybrid approach to resolve repetitive regions.

Improvements are expected not only in sequencing platforms but also in assemblers. Recently, two groups assessed the quality of the currently available assemblers. The 2011 Assemblathon was the first competition among assemblers [65]. For this competition, simulated data were generated and groups of assemblers were asked to blindly assemble it. The use of simulated data poses a problem in determining the applicability of the results to other data sets. The 2012 GAGE (Genome Assembly Gold-Standard Evaluations) competition for assembling real data resulted in the following conclusions: (1) the data quality has a greater influence on the final outcome

than the assembler itself; and (2) the results do not support the current measures of correctness (related to contiguity)[26].

There is a large gap between the availability of genomic sequences in databases and the commercial production of vaccines and antibiotics in recent years, especially in the fields of investment and success (“expected return”). Drug development for all potential targets and effective vaccines has produced limited success. In contrast, there has been an acceleration in the discovery of new targets due to the refinement of bioinformatics tools for this purpose, such as epitope mapping and searching for secreted proteins. However, the major problems facing vaccine and antibiotic development, such as resistance mechanisms and host immune responses, remain unsolved.

Genome analysis constitutes a strategy for the expansion and diversification of the pharmacology and vaccinology sectors. This methodology can be used to explore a large number of targets and to reduce the costs of molecular and immunological tests. Finally, to improve the production of antibiotics and vaccines, it is necessary to know more about bacterial regulatory pathways. New interactome and microbiome studies must be implemented to assist this search.

### **Acknowledgements**

This work involved the collaboration of various institutions, including the Genomics and Proteomics Network of the State of Pará of the Federal University of Pará (Rede Paraense de Genômica e Proteômica da Universidade Federal do Pará), the Amazon Research Foundation (Fundação Amazônia Paraense - FAPESPA), the National Council for Scientific and Technological Development (Conselho Nacional de Desenvolvimento Científico e Tecnológico -CNPq), the Brazilian Federal Agency for the Support and Evaluation of Graduate Education (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior -CAPES) and the Minas Gerais Research Foundation (Fundação de Amparo à Pesquisa do estado de Minas Gerais).

### **Footnotes**

Supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) in Brazil, processes BEX 12954-12-8 and 11517-12-3, to Barbosa EGV and Aburjaile FF.

P- Reviewers: Bhattacharya SK, Faik A S- Editor: Ma YJ L- Editor: A E- Editor: Lu YJ

## References

1. Zhang J, Chiodini R, Badr A, Zhang G. The impact of next-generation sequencing on genomics. *J Genet Genomics*. 2011;38:95–109. [[PMC free article](#)] [[PubMed](#)]
2. Prachi P, Donati C, Masciopinto F, Rappuoli R, Bagnoli F. Deep sequencing in pre- and clinical vaccine research. *Public Health Genomics*. 2013;16:62–68. [[PubMed](#)]
3. Pucci MJ. Use of genomics to select antibacterial targets. *Biochem Pharmacol*. 2006;71:1066–1072. [[PubMed](#)]
4. Mills SD. When will the genomics investment pay off for antibacterial discovery? *Biochem Pharmacol*. 2006;71:1096–1102. [[PubMed](#)]
5. Pareek CS, Smoczynski R, Tretyn A. Sequencing technologies and genome sequencing. *J Appl Genet*. 2011;52:413–435. [[PMC free article](#)] [[PubMed](#)]
6. Liu L, Li Y, Li S, Hu N, He Y, Pong R, Lin D, Lu L, Law M. Comparison of next-generation sequencing systems. *J Biomed Biotechnol*. 2012;2012:251364. [[PMC free article](#)] [[PubMed](#)]
7. Metzker ML. Sequencing technologies-the next generation. *Nat Rev Genet*. 2010;11:31–46. [[PubMed](#)]
8. Magi A, Benelli M, Gozzini A, Girolami F, Torricelli F, Brandi ML. Bioinformatics for Next Generation Sequencing Data. *Genes*. 2010;1:294–307. [[PMC free article](#)] [[PubMed](#)]
9. Loman NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, Pallen MJ. Performance comparison of benchtop high-throughput sequencing platforms. *Nat Biotechnol*. 2012;30:434–439. [[PubMed](#)]
10. Klassen JL, Currie CR. Gene fragmentation in bacterial draft genomes: extent, consequences and mitigation. *BMC Genomics*. 2012;13:14. [[PMC free article](#)] [[PubMed](#)]
11. Miller JR, Koren S, Sutton G. Assembly algorithms for next-generation sequencing data. *Genomics*. 2010;95:315–327. [[PMC free article](#)] [[PubMed](#)]
12. Bashir A, Klammer AA, Robins WP, Chin CS, Webster D, Paxinos E, Hsu D, Ashby M, Wang S, Peluso P, et al. A hybrid approach for the automated finishing of bacterial genomes. *Nat Biotechnol*. 2012;30:701–707. [[PMC free article](#)] [[PubMed](#)]
13. Ribeiro FJ, Przybylski D, Yin S, Sharpe T, Gnerre S, Abouelleil A, Berlin AM, Montmayeur A, Shea TP, Walker BJ, et al. Finished bacterial genomes from shotgun sequence data. *Genome Res*. 2012;22:2270–2277. [[PMC free article](#)] [[PubMed](#)]

14. Baker M. De novo genome assembly: what every biologist should know. *Nature Methods*. 2012;9:333–337.
15. Salzberg SL, Phillippy AM, Zimin A, Puiu D, Magoc T, Koren S, Treangen TJ, Schatz MC, Delcher AL, Roberts M, et al. GAGE: A critical evaluation of genome assemblies and assembly algorithms. *Genome Res*. 2012;22:557–567. [[PMC free article](#)] [[PubMed](#)]
16. Faircloth BC, Glenn TC. Not all sequence tags are created equal: designing and validating sequence identification tags robust to indels. *PLoS One*. 2012;7:e42543. [[PMC free article](#)] [[PubMed](#)]
17. Diguistini S, Liao NY, Platt D, Robertson G, Seidel M, Chan SK, Docking TR, Birol I, Holt RA, Hirst M, et al. De novo genome sequence assembly of a filamentous fungus using Sanger, 454 and Illumina sequence data. *Genome Biol*. 2009;10:R94. [[PMC free article](#)] [[PubMed](#)]
18. Chain PS, Grafham DV, Fulton RS, Fitzgerald MG, Hostetler J, Muzny D, Ali J, Birren B, Bruce DC, Buhay C, et al. Genomics. Genome project standards in a new era of sequencing. *Science*. 2009;326:236–237. [[PMC free article](#)] [[PubMed](#)]
19. Shendure J, Mitra RD, Varma C, Church GM. Advanced sequencing technologies: methods and goals. *Nat Rev Genet*. 2004;5:335–344. [[PubMed](#)]
20. Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM, Wang MD, Zhang K, Mitra RD, Church GM. Accurate multiplex polony sequencing of an evolved bacterial genome. *Science*. 2005;309:1728–1732. [[PubMed](#)]
21. Richardson P. Special Issue: Next Generation DNA Sequencing. *Genes*. 2010;1:385–387.
22. Munroe DJ, Harris TJ. Third-generation sequencing fireworks at Marco Island. *Nat Biotechnol*. 2010;28:426–428. [[PubMed](#)]
23. Kececioğlu J, Ju J. Separating repeats in DNA sequence assembly. *Proceedings of the 5th International Conference on Computational Biology, 2001*. New York: ACM; 2001. pp. 176–183.
24. Phillippy AM, Schatz MC, Pop M. Genome assembly forensics: finding the elusive mis-assembly. *Genome Biol*. 2008;9:R55. [[PMC free article](#)] [[PubMed](#)]
25. Soares SC, Abreu VA, Ramos RT, Cerdeira L, Silva A, Baumbach J, Trost E, Tauch A, Hirata R, Mattos-Guaraldi AL, et al. PIPS: pathogenicity island prediction software. *PLoS One*. 2012;7:e30848. [[PMC free article](#)] [[PubMed](#)]

26. Ricker N, Qian H, Fulthorpe RR. The limitations of draft assemblies for understanding prokaryotic adaptation and evolution. *Genomics*. 2012;100:167–175. [[PubMed](#)]
27. Pop M. Genome assembly reborn: recent computational challenges. *Brief Bioinform*. 2009;10:354–366. [[PMC free article](#)] [[PubMed](#)]
28. Tsai IJ, Otto TD, Berriman M. Improving draft assemblies by iterative mapping and assembly of short reads to eliminate gaps. *Genome Biol*. 2010;11:R41. [[PMC free article](#)] [[PubMed](#)]
29. Cerdeira LT, Carneiro AR, Ramos RT, de Almeida SS, D’Afonseca V, Schneider MP, Baumbach J, Tauch A, McCulloch JA, Azevedo VA, et al. Rapid hybrid de novo assembly of a microbial genome using only short reads: *Corynebacterium pseudotuberculosis* I19 as a case study. *J Microbiol Methods*. 2011;86:218–223. [[PubMed](#)]
30. Boetzer M, Pirovano W. Toward almost closed genomes with GapFiller. *Genome Biol*. 2012;13:R56. [[PMC free article](#)] [[PubMed](#)]
31. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA*. 1977;74:5463–5467. [[PMC free article](#)] [[PubMed](#)]
32. de Boer HA, Gilbert SF, Nomura M. DNA sequences of promoter regions for rRNA operons *rrnE* and *rrnA* in *E. coli*. *Cell*. 1979;17:201–209. [[PubMed](#)]
33. Nakamura K, Inouye M. DNA sequence of the gene for the outer membrane lipoprotein of *E. coli*: an extremely AT-rich promoter. *Cell*. 1979;18:1109–1117. [[PubMed](#)]
34. Porter AG, Barber C, Carey NH, Hallewell RA, Threlfall G, Emtage JS. Complete nucleotide sequence of an influenza virus haemagglutinin gene from cloned DNA. *Nature*. 1979;282:471–477. [[PubMed](#)]
35. Messing J, Crea R, Seeburg PH. A system for shotgun DNA sequencing. *Nucleic Acids Res*. 1981;9:309–321. [[PMC free article](#)] [[PubMed](#)]
36. Brown NL, Misra TK, Winnie JN, Schmidt A, Seiff M, Silver S. The nucleotide sequence of the mercuric resistance operons of plasmid R100 and transposon Tn501: further evidence for mer genes which enhance the activity of the mercuric ion detoxification system. *Mol Gen Genet*. 1986;202:143–151. [[PubMed](#)]
37. Postle K, Good RF. DNA sequence of the *Escherichia coli* *tonB* gene. *Proc Natl Acad Sci USA*. 1983;80:5235–5239. [[PMC free article](#)] [[PubMed](#)]

38. Overduin P, Boos W, Tommassen J. Nucleotide sequence of the *ugp* genes of *Escherichia coli* K-12: homology to the maltose system. *Mol Microbiol.* 1988;2:767–775. [[PubMed](#)]
39. Sutton GG, White O, Adams MD and Kerlavage A. TIGR Assembler: A new tool for assembling large shotgun sequencing projects. *Genome Sci Technol.* 1995;1 Suppl 1:S9–S19.
40. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science.* 1995;269:496–512. [[PubMed](#)]
41. Fraser CM, Eisen JA, Nelson KE, Paulsen IT, Salzberg SL. The value of complete microbial genome sequencing (you get what you pay for) *J Bacteriol.* 2002;184:6403–6405; discussion 6405. [[PMC free article](#)][[PubMed](#)]
42. Smith DR. Death of the genome paper. *Front Genet.* 2013;4:72. [[PMC free article](#)] [[PubMed](#)]
43. Medini D, Donati C, Tettelin H, Massignani V, Rappuoli R. The microbial pan-genome. *Curr Opin Genet Dev.* 2005;15:589–594. [[PubMed](#)]
44. Soares SC, Silva A, Trost E, Blom J, Ramos R, Carneiro A, Ali A, Santos AR, Pinto AC, Diniz C, et al. The pan-genome of the animal pathogen *Corynebacterium pseudotuberculosis* reveals differences in genome plasticity between the biovar *ovis* and *equi* strains. *PLoS One.* 2013;8:e53818. [[PMC free article](#)] [[PubMed](#)]
45. Jakobsen TH, Hansen MA, Jensen PØ, Hansen L, Riber L, Cockburn A, Kolpen M, Rønne Hansen C, Ridderberg W, Eickhardt S, et al. Complete genome sequence of the cystic fibrosis pathogen *Achromobacter xylosoxidans* NH44784-1996 complies with important pathogenic phenotypes. *PLoS One.* 2013;8:e68484. [[PMC free article](#)] [[PubMed](#)]
46. Parkhill J. In defense of complete genomes. *Nat Biotechnol.* 2000;18:493–494. [[PubMed](#)]
47. Branscomb E, Predki P. On the high value of low standards. *J Bacteriol.* 2002;184:6406–6409; discussion 6409. [[PMC free article](#)] [[PubMed](#)]
48. Kingsford C, Schatz MC, Pop M. Assembly complexity of prokaryotic genomes using short reads. *BMC Bioinformatics.* 2010;11:21. [[PMC free article](#)] [[PubMed](#)]
49. Dagan T. Phylogenomic networks. *Trends Microbiol.* 2011;19:483–491. [[PubMed](#)]
50. Schnoes AM, Brown SD, Dodevski I, Babbitt PC. Annotation error in public databases: misannotation of molecular function in enzyme superfamilies. *PLoS Comput Biol.* 2009;5:e1000605. [[PMC free article](#)][[PubMed](#)]



51. Rappuoli R. Reverse vaccinology. *Curr Opin Microbiol.* 2000;3:445–450. [[PubMed](#)]
52. Wooldridge K. Bacterial secreted proteins: secretory mechanisms and role in pathogenesis. United States: Caister Academic Press; 2009. pp. 300–315.
53. Simeone R, Bottai D, Brosch R. ESX/type VII secretion systems and their role in host-pathogen interaction. *Curr Opin Microbiol.* 2009;12:4–10. [[PubMed](#)]
54. Stavrinides J, McCann HC, Guttman DS. Host-pathogen interplay and the evolution of bacterial effectors. *Cell Microbiol.* 2008;10:285–292. [[PubMed](#)]
55. He Y, Xiang Z, Mobley HL. Vaxign: the first web-based vaccine design program for reverse vaccinology and applications for vaccine development. *J Biomed Biotechnol.* 2010;2010:297505. [[PMC free article](#)] [[PubMed](#)]
56. Santos AR, Pereira VB, Barbosa E, Baumbach J, Pauling J, Röttger R, Turk MZ, Silva A, Miyoshi A, Azevedo V. Mature Epitope Density--a strategy for target selection based on immunoinformatics and exported prokaryotic proteins. *BMC Genomics.* 2013;14 Suppl 6:S4. [[PMC free article](#)] [[PubMed](#)]
57. Seib KL, Zhao X, Rappuoli R. Developing vaccines in the era of genomics: a decade of reverse vaccinology. *Clin Microbiol Infect.* 2012;18 Suppl 5:109–116. [[PubMed](#)]
58. Tettelin H. The bacterial pan-genome and reverse vaccinology. *Genome Dyn.* 2009;6:35–47. [[PubMed](#)]
59. Donati C, Rappuoli R. Reverse vaccinology in the 21st century: improvements over the original design. *Ann N Y Acad Sci.* 2013;1285:115–132. [[PubMed](#)]
60. Walsh C. Where will new antibiotics come from? *Nat Rev Microbiol.* 2003;1:65–70. [[PubMed](#)]
61. Silver LL. Challenges of antibacterial discovery. *Clin Microbiol Rev.* 2011;24:71–109. [[PMC free article](#)] [[PubMed](#)]
62. Payne DJ, Gwynn MN, Holmes DJ, Pompliano DL. Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat Rev Drug Discov.* 2007;6:29–40. [[PubMed](#)]
63. Projan SJ. Why is big Pharma getting out of antibacterial drug discovery? *Curr Opin Microbiol.* 2003;6:427–430. [[PubMed](#)]
64. Bush K, Pucci MJ. New antimicrobial agents on the horizon. *Biochem Pharmacol.* 2011;82:1528–1539. [[PubMed](#)]

65. Earl D, Bradnam K, St John J, Darling A, Lin D, Fass J, Yu HO, Buffalo V, Zerbino DR, Diekhans M, et al. Assemblathon 1: a competitive assessment of de novo short read assembly methods. *Genome Res.*2011;21:2224–2241. [[PMC free article](#)] [[PubMed](#)]

#### **Anexo 4: Nova montagem e anotação do genoma de *Propionibacterium freudenreichii* CIRM-BIA138**

Com o intuito de finalizar o genoma alvo da tese, referente a linhagem de *Propionibacterium freudenreichii* CIRM 138, foi proposto como parte deste trabalho uma nova montagem e anotação deste genoma. Para isso, diferentes abordagens foram realizadas na etapa de montagem: a primeira tentativa foi realizada com a associação do programa *CLC Workbench 7.0* com *CONTIGuator 2.7* (*scaffolding* por referência com a CIRM-BIA1). Posteriormente ao alinhamento dos dados brutos gerou-se um *scaffold* e estes foram alinhados contra o genoma de referência (CIRM-BIA1). O fechamento manual de *gaps* foi realizado com o auxílio do programa *CLC Workbench 7.0* e pelo banco de dados do *Basic Local Alignment Search Tool (BLAST)*, e, finalmente, 28 contigs foram obtidos como resultado final deste genoma. Outras abordagens de montagem de genomas, tais que o *SPAdes 3.5* e *scaffolding de novo* com o *SSPACE 2*, também foram testadas com o objetivo de finalizá-lo, porém não apresentaram resultados promissores.

A anotação deste genoma foi realizada em duas etapas: automática e manual. A anotação automática foi executada pelo *RAST Server (Rapid Annotations using Subsystems Technology* - <http://rast.nmpdr.org/rast.cgi>) e a anotação manual foi realizada através do programa *Artemis - Genome Browser and Annotation Tool*, a partir do arquivo .embl. Neste arquivo, foram verificadas as posições corretas dos códons iniciais e de parada, além da curadoria de todas as sequências codificantes: CDS (*Coding Sequences*), por meio da combinação de diversos bancos de dados, como: *BLAST* (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), *Uniprot* (<http://www.uniprot.org>), *PFAM* (<http://pfam.xfam.org/>), *InterProScan 5* (<http://www.ebi.ac.uk/Tools/pfa/iprscan5/>) e *transporter classification database* (<http://www.tcdb.org/>).

#### **Annexe 4 : Nouvel assemblage et annotation du génome de *Propionibacterium freudenreichii* CIRM-BIA138**

Afin de finaliser le génome cible de la thèse relative à la souche de *Propionibacterium freudenreichii* CIRM-BIA138, ont été proposé dans le cadre de ce travail, une nouvelle assemblage et annotation de ce génome. A cet effet, différentes approches ont été réalisés dans l'étape de l'assemblage: la première tentative a été réalisée avec l'association du logiciel *CLC Workbench 7.0* com *CONTIGuator 2.7* (*scaffolding* par référence avec la souche de *Propionibacterium freudenreichii* CIRM-BIA1), ensuite l'alignement des données brutes générés un *scaffold* et aussi ils ont été obtenus contre le génome de référence (CIRM-BIA1). La clôture manuelle des *gaps* a été effectuée avec l'aide d'un logiciel *CLC Workbench 7.0* et avec la base de données du *Basic Local Alignment Search Tool (BLAST)*, et enfin 28 contigs ont été obtenues comme le résultat final de ce génome. D'autres approches ont été testés avec l'objectif de leur finaliser, comme le *SPAdes 3.5* et *scaffolding de novo* avec le *SSPACE 2*, mais ne montrent pas des résultats prometteurs.

L'annotation de ce génome a été réalisée en deux étapes : automatique et manuelle. L'annotation automatique a été effectuée par *RAST Server (Rapid Annotations using Subsystems Technology* - <http://rast.nmpdr.org/rast.cgi>) et l'annotation manuelle a été réalisée en utilisant le programme *Artemis - Genome Browser and Annotation Tool*, à partir de l'archive .embl. Dans ce fichier, les positions correctes des codons initiales et des codons d'arrêt ont été trouvés, ainsi que la vérification du génome des toutes les séquences codantes : CDS (*Coding Sequences*) en combinant plusieurs bases de données, tels que *PFAM* (<http://pfam.xfam.org/>), *InterProScan 5* (<http://www.ebi.ac.uk/Tools/pfa/iprscan5/>) et *transporter classification database* (<http://www.tcdb.org/http://www.tcdb.org/>).

## Anexo 5: Dados genômicos do *Genomes Online Database (GOLD)*

Projetos distribuídos de acordo com o seu domínio

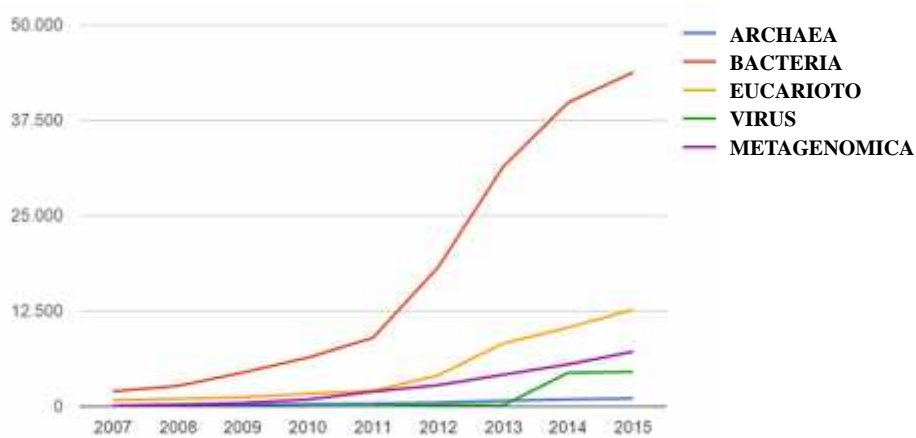


Figura 13: Número total de projetos por ano depositados no GOLD (<https://gold.jgi.doe.gov/statistics>).

Distribuição de projetos genômicos em bactérias



Figura 14: Distribuição de projetos genômicos em bactérias de acordo com sua classificação (<https://gold.jgi-psf.org/distribution>).

Relevância dos projetos bacterianos



Figura 15: Relevância dos projetos bacterianos divididos de acordo com sua categoria (disponível em <https://gold.jgi-psf.org/statistics> acesso em 11/06/15).

### Filogenia dos projetos bacterianos

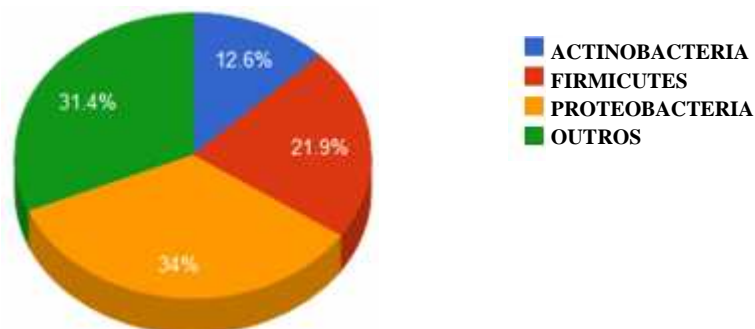


Figura 16: Projetos bacterianos organizados de acordo com sua filogenia (disponível em <https://gold.jgi-psf.org/statistics> acesso em 11/06/15).

### Depósito de genomas no GOLD segundo o gênero bacteriano

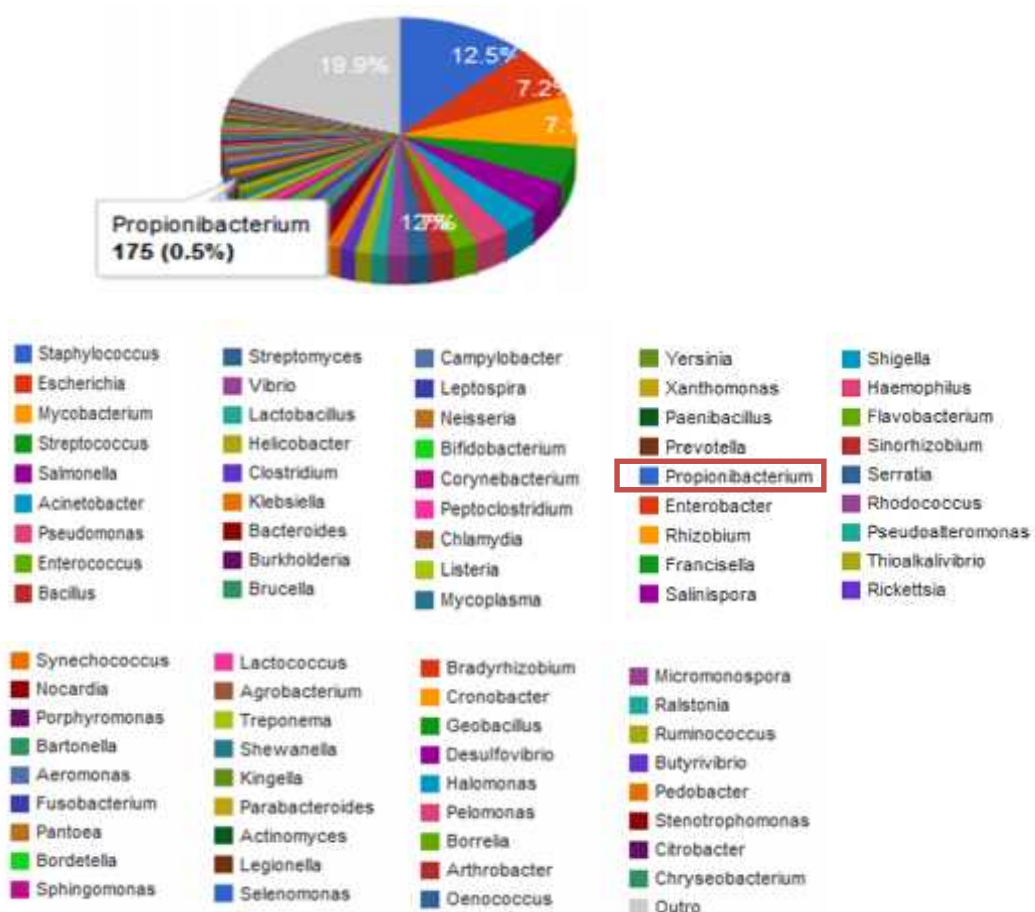


Figura 17: Porcentagem de projetos depositados no GOLD de acordo com seu gênero bacteriano (disponível em <https://gold.jgi-psf.org/distribution> acesso em 11/06/15).

## Annexe 5 : Données génomiques du *Genomes Online Database (GOLD)*

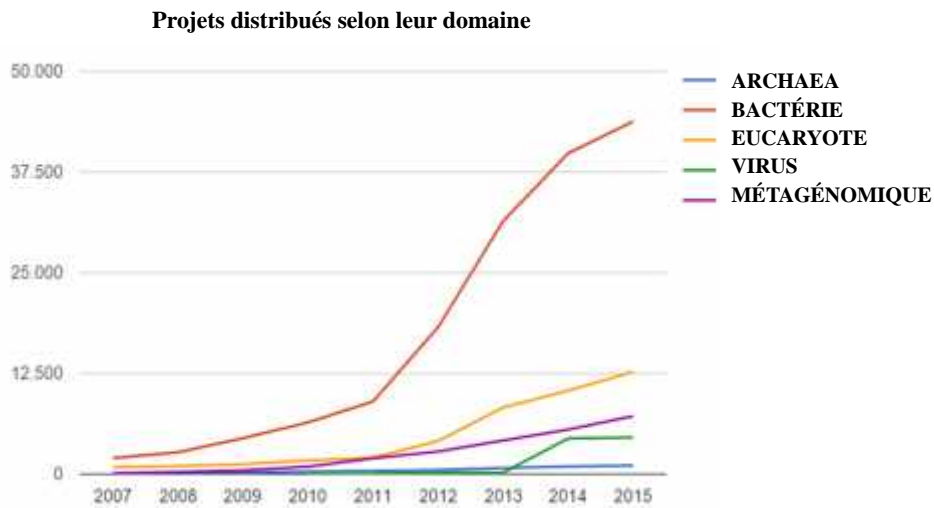


Figure 13 : Nombre total de projets par an déposés dans le GOLD (<https://gold.jgi-psf.org/distribution>).

### Distribution des projets génomiques dans les bactéries

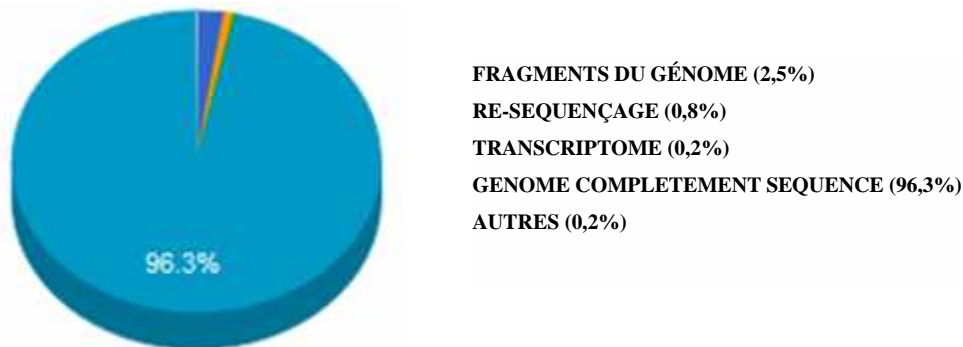


Figure 14 : Distribution des projets génomiques dans les bactéries selon leur classement (<https://gold.jgi-psf.org/distribution>).

### Pertinence des projets bactériens

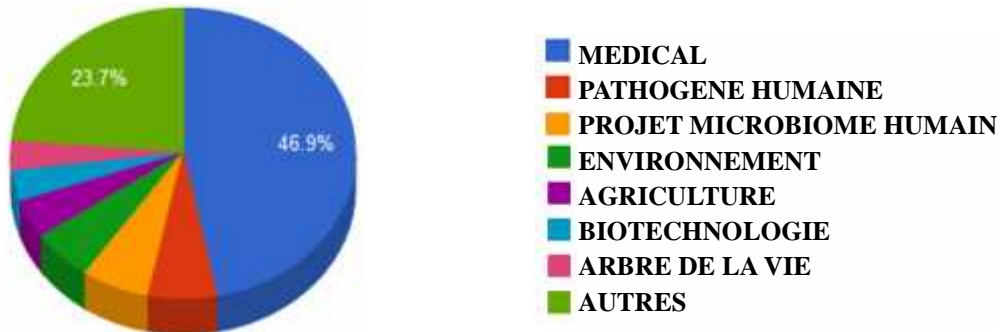


Figure 15: La pertinence des projets bactériens présentés selon leur catégorie (disponible en <https://gold.jgi-psf.org/statistics> accès 06.11.15).

### Phylogénie des projets bactériens

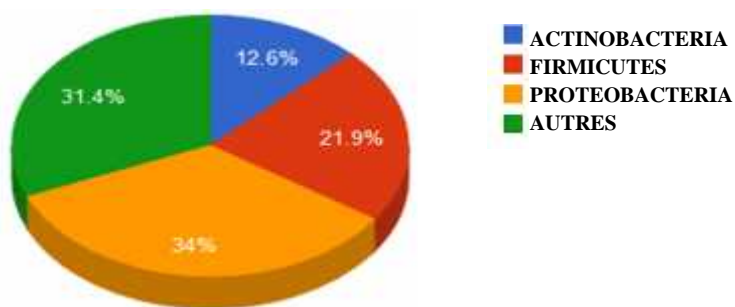


Figure 16 : Projets bactériens organisé selon leur phylogénie (disponible en <https://gold.jgi-psf.org/statistics> accès 06.11.15).

### Projets déposés dans GOLD selon leur genre bactérien

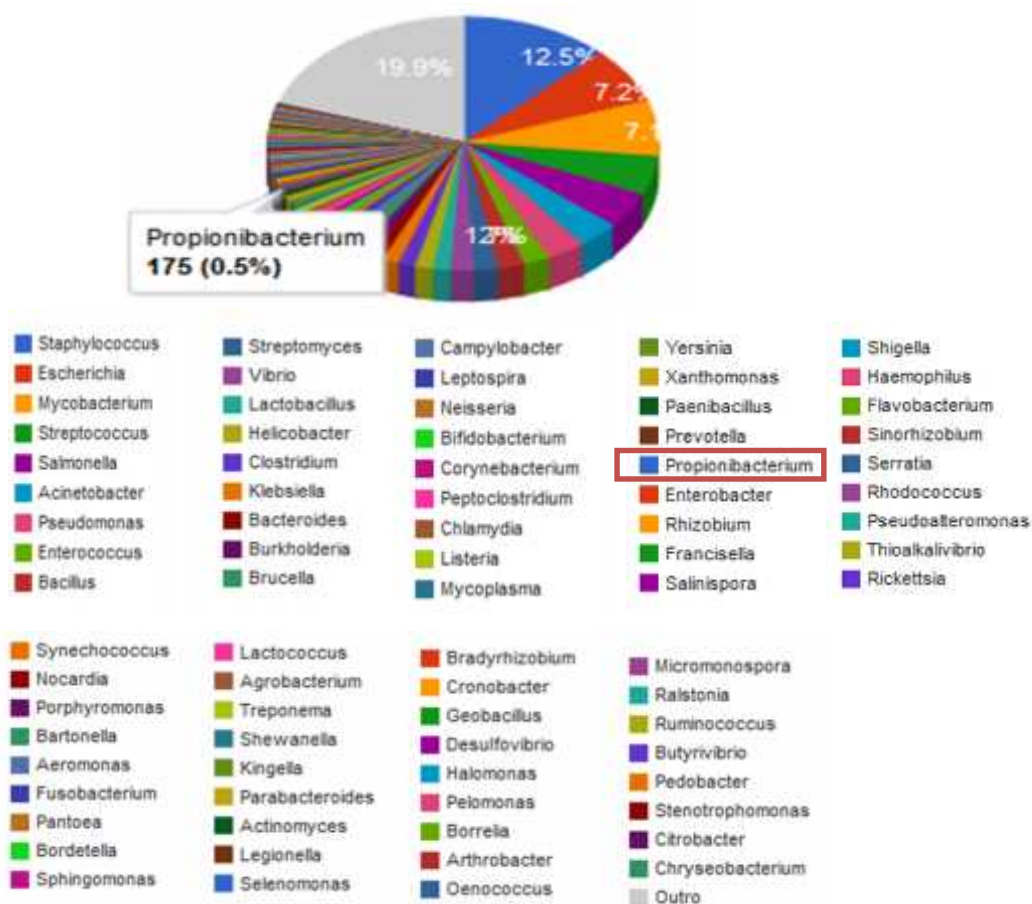


Figure 17 : Pourcentage des projets déposés dans GOLD selon leur genre bactérien (disponible en <https://gold.jgi-psf.org/distribution> accès en 11/6/15).



## **Anexo 6: Manuscrito em redação - Survival strategies of Gram-positive bacteria**

## **Annexe 6 : Manuscrit en cours - Survival strategies of Gram-positive bacteria**

### **Survival strategies of Gram-positive bacteria (Review)**

Flavia Figueira Aburjaile<sup>1,2,3</sup>, H el ene Falentin<sup>2,3</sup>, Anderson Miyoshi<sup>1</sup>, Yves Le Loir<sup>2,3</sup>, Vasco Azevedo<sup>1</sup>

<sup>1</sup> Department of General Biology, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

<sup>2</sup> INRA, UMR 1253, Science et Technologie du Lait et de l' euf, 35000, Rennes, France

<sup>3</sup> AGROCAMPUS OUEST, UMR1253, UMR Science et Technologie du Lait et de l' euf, 35000, Rennes, France

#### **Abstract**

Bacteria have been subject of studies, since antiquity, can cause disease to humans, animals and plants, and they can be used for therapeutic and biotechnology applications due to their beneficial properties, such as probiotics. Due to a multitude of information that covers this topic, text mining methods that allow automated extraction of information for the investigation of these organisms have become essential. In this review the text-mining with the LitProf program allowed the selection of a well labored content by combining biological concepts and growth phases and strategies of gram-positive bacteria that allow their persistence and viability for longer periods. Further we will discuss the criteria for the classification of bacterial growth phases and the expression of various phenotypes, while passing through a series of physiological and morphological cellular adjustments aimed at maintaining the survival of these populations. Therefore, regulation of bacterial metabolism is needed, for example, nutrient storage. Furthermore, strategies that allow bacterial survival to various stresses, such as programmed cell death and dormancy will be addressed in this discussion. Bacteria that utilize dormancy are also capable of maintaining viable but not culturable in unfavorable situations can be revived and set their conditions for bacterial growth when exposed to favorable conditions again. In the future, the manipulation and regulation of bacterial strategies will allow technological advances in diverse industrial sectors such as agrifood, pharmaceutical, cosmetic, biomedical and forestry.

**Keywords:** bacteria, Gram-positive, long-term survival, survival strategies.

## **Introduction**

Bacteria are ubiquitous microorganisms, known for their ability to multiply and adapt in many environments and growth conditions (Roszak, DB, e Colwell, RR, de 1987; Tempest e Wouters, 1981; M. Neijssel, e W. Zevenboom, 1983). For decades, bacterial survival has been a subject of great interest in various fields of microbiology and these studies crosslinked with many other disciplines such as ecology, genetics, bioinformatics, biotechnology, food, epidemiology or infectiology. A better understanding of bacterial survival can be helpful in human, animal or plant health issues as well as in the biotechnology or food industry. With regard to survival capacities, bacteria can indeed be divided into two groups: (1) pathogenic bacteria, that are able to cause infections, and whose threat prompt the need for decontamination procedures, vaccines and drugs and (2) non-pathogenic bacteria, some of which are useful for industrial purpose like the production of biomolecules or the development of probiotic products, which in some applications require a good survival to maintain activity.

Bacterial growth comprises different phases (lag, exponential, stationary, and death) along which the bacterial population copes with continuously changing conditions with, among others, high and low availability of nutrients (Helmus et al., 2012). Bacteria are thus able to switch on or off or to modulate the expression of many mechanisms to ensure their homeostasis.

To overcome harsh conditions, some bacterial species sporulate. This strategy has been widely documented and reviewed in detailed elsewhere (Palop, MañAs and CondóN 1999). Besides, some non-sporulating species are nevertheless able to survive for long periods of time and their growth enters a phase called “long-term survival” (LTS). In 1998, Watson and collaborators defined ‘long-term survival’ as a period of survival in which bacteria can respond to the stimulus of hunger and remain alive and viable. For other authors, this phase of survival is characterized, in addition to starvation, by a resistance to various stresses (Wen et al. 2011). According to Finkel (Finkel 2006a), this is a period in which the bacteria undergo competition for nutrients and most likely have reduced metabolic capacity and exhibit a particular phenotype. During the LTS phase, many phenotypic changes can indeed be observed: a reduced metabolism, biofilm formation, changes in shape and size of bacterial cells, changes in size and color of colonies, and or viable but nonculturable phenotype. These changes are part of various strategies bacterial species have evolved to ensure a longer viability.

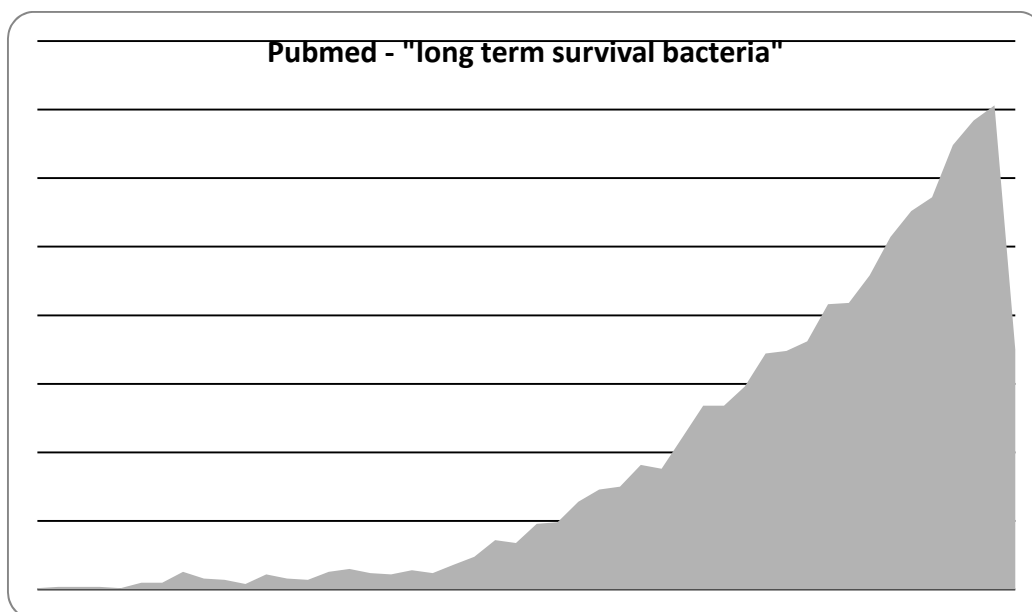
Depending on the study or the species considered, LTS spans a range of time that goes from hours to years (Zambrano and Kolter 1993; Finkel 2006a).

Several strategies bacteria can be used to survive for long periods, for example by using different substrates available in the medium, followed by the ability to enter a resting state for long periods (with low metabolic activity) or by entering a reversible cell stasis, after which these bacteria can resume cellular multiplication when conditions are favorable (Greenblatt et al. 2004b).

In this review article, we will address the strategies Gram-positive bacteria exert during LTS periods, aiming at their metabolic maintenance and proliferation capacity. Sporulation will not be considered here.

### **Evolution of research dealing with long-term survival**

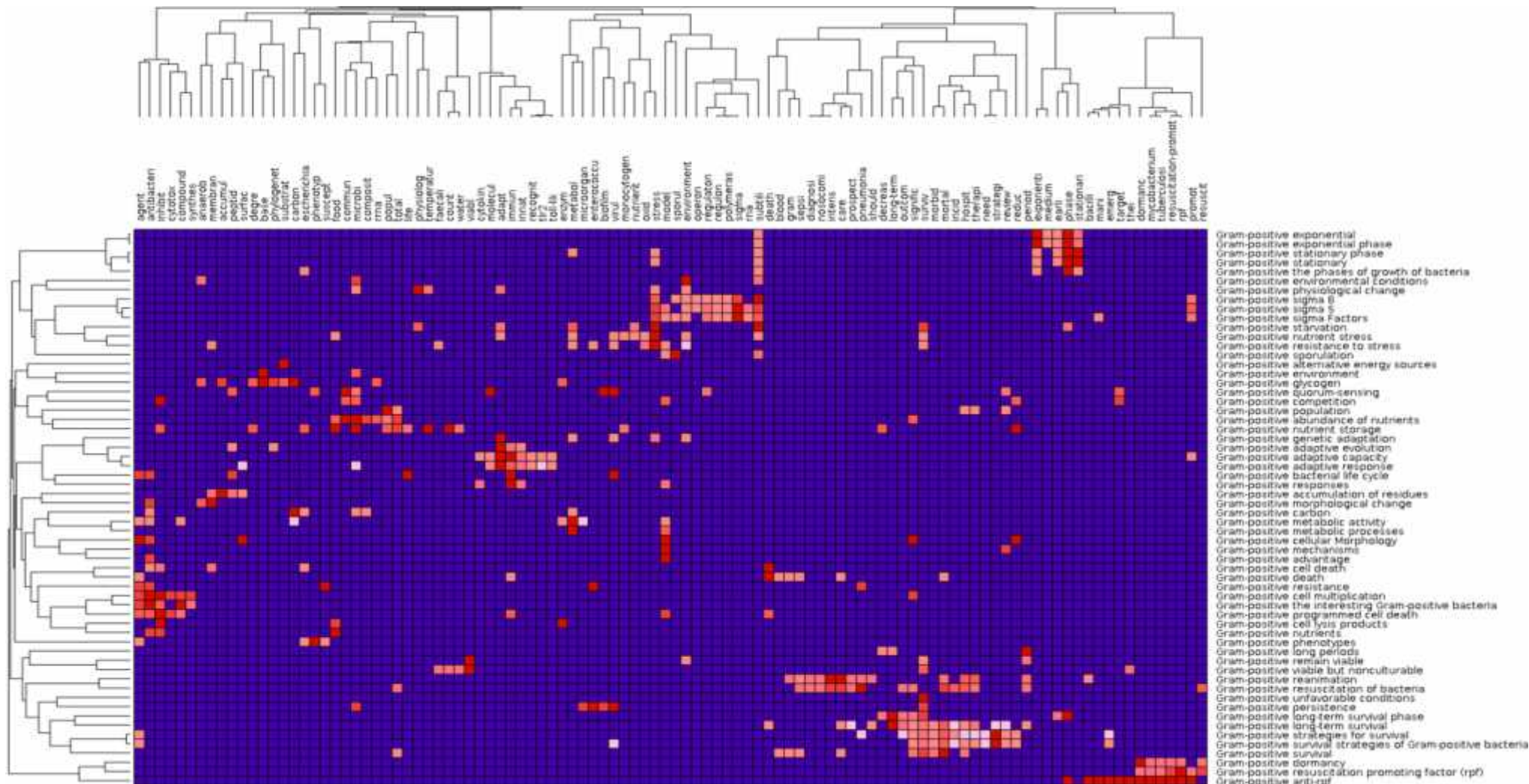
Research works dated from 1967 mark the beginning of the scientific community's interest for bacterial survival in extended periods. Since then, a gradual increase in publications can be observed with this issue in PubMed (<http://www.ncbi.nlm.nih.gov/pubmed>) (Figure 1).



**Figure 1:** The data in PubMed results by year in 01/09/2015 (<http://www.ncbi.nlm.nih.gov/pubmed/?term=long-term+survival+bacteria>).

Among the hundreds of references found in databases, the terminology used to describe LTS is quite variable. A short text mining study applied to the terms associated to LTS was carried out using the LitProf program (Chassaubel and Sher, 2002). The literature profiles were generated by extracting the frequencies of Gram positive- and

LTS-related terms from hundreds of abstracts stored in the PubMed database. The terms were filtered on the basis of both repetitive occurrence and co-occurrence and clustered with regard to their frequency values. This resulted in a picture of the functional relationship among the lists of LTS-related terms (Figure 2). Data analysis revealed strong relationship between the LTS-related terms and the terms associated to the description of growth phases, dormancy, resuscitation, etc. It also clearly shows a link with the ability of bacteria to trigger a series of mechanisms involved in the adaptation in response to adverse environmental factors. Interestingly, the results show associations between bacterial virulence and persistence, which is relevant, e.g. in nosocomial infections, and between cell-lysis products and food, which is relevant, e.g. the involvement of lactic acid bacteria in chesse ripening. Most the associations revealed by this data analysis will be presented in the following paragraphs.



1  
 2 Figure 2: Bow's default parameters, except for  $F = 0.5$ , and hierarchical clustering with Pearson correlation and UPGMA to group X (93terms) and Y (60 queries)  
 3 axes.

## The phases of bacterial growth

The interest in the investigation and characterization of bacterial growth phases began for almost 100 years. Since 1917, in accordance with fundamental theories in physics and mathematical theories the researchers suggested that bacteria have a growth cycle composed of 7 steps: initial stationary phase, lag phase (or positive growth acceleration phase), logarithmic growth phase, negative growth acceleration, maximum stationary phase, phase of accelerated death and logarithm death phase (Buchanan 1918). Later on, Lane-Clayton simplified the description of bacterial phases to four divisions: initial period of slow growth (or even no growth), period of regular growth, period during which numbers remain stable and period during which the numbers of living bacteria are decreasing (Buchanan 1918). In 2006, Finkel divided the bacterial growth into five phases: lag, exponential (or logarithmic), stationary, death and long-term stationary phase (Finkel 2006a). These final steps will be presented in more detail in the Table 1.

**Table 1: Description the phases of bacterial growth cycle.**

<b>Phases of bacterial growth</b>	<b>Description</b>
<b>Lag</b>	First stage, which consists in the adaptation of the bacteria to the environment in which it is exposed. In this phase, bacteria adopt a high metabolic rate, and consequently a high protein synthesis for cell maintenance.
<b>Exponential</b>	Second stage. Micro-organisms tend to consume all the nutrients in the medium to generate an energy reserve, with high activity of the DNA polymerase III, involved in DNA replication.
<b>Stationary</b>	Third stage, the processes that lead to cell death will begin to appear in some bacterial species. The oxidative stress induces the bacteria to activate a number of mechanisms to maintain viability.
<b>Death</b>	Stimuli such as a lack of available nutrients or the presence of toxic metabolites in the medium can induce metabolic activities causing cell death in bacteria (unable to multiply).
<b>Long-Term-Survival</b>	The long-term survival phase presents characteristics that associated with bacterial persistence, including morphological alterations. Bacteria adapt to the environment, and trigger mechanisms to ensure their survival.

In Figure 3, we briefly present the five important bacterial phases: (1) lag phase, (2) exponential phase, (3) stationary phase, (4) death phase and (5) long-term survival.

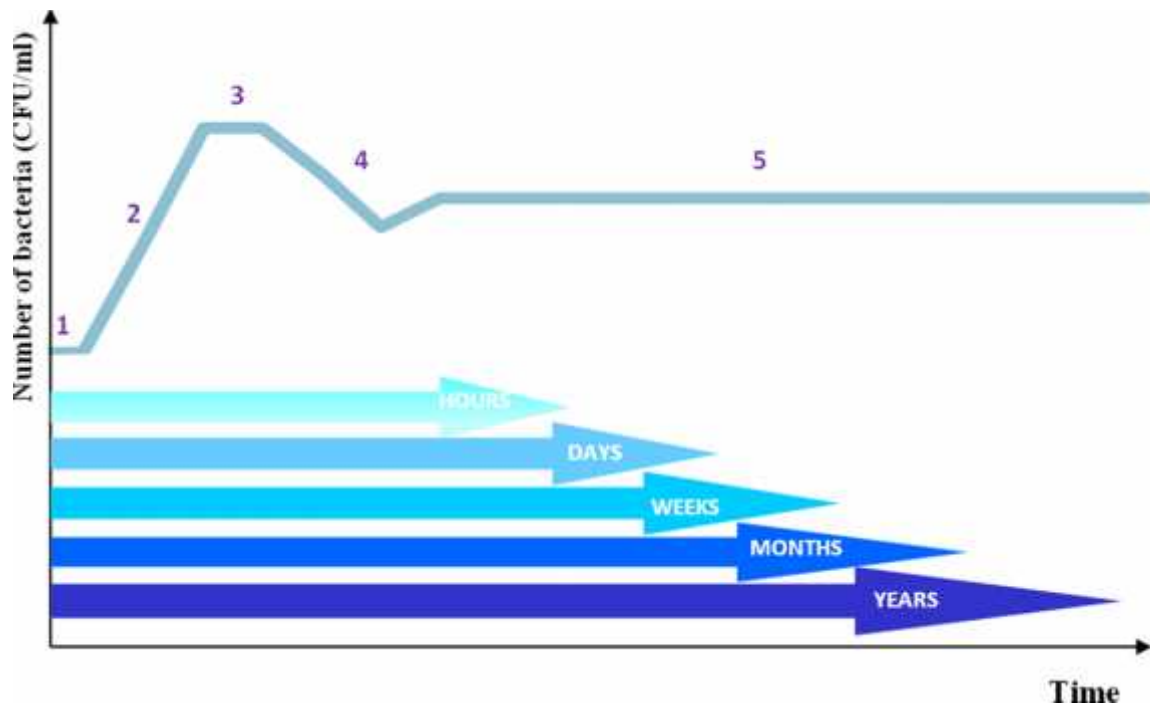


Figure 3: Bacterial growth shows about the same phases but the time scale ranges from hours to years depending on the species.

### Bacterial phenotypes in long term survival

#### Cellular and colony morphological changes

During the LTS, Gram positive bacteria can undergo morphological changes, at the cellular or the colony level (Table 2).

In bacteria, nutrient availability has a great impact on bacterial cell size (Chien et al., 2012). Under stress conditions such as starvation encountered in LTS phase, some Gram positive bacteria can enter into a cell-wall deficient state also known as L-form conversion. This phenomenon is well-described in pathogenic bacteria of Mycobacterium genus. In *M. bovis* (Slavchev; Michailoca; Markova, 2013), or *M. tuberculosis* (Markova et al., 2012), L-form conversion was shown to be an adaptive strategy to survive and reproduce under unfavorable conditions. In other Gram positive species, a reduction in cell size is observed when the population enters the LTS phase. This was shown in *S. aureus* which indeed harbours a decrease in cell size in starvation survival. This phenotype is accompanied by an increase in resistance to acid shock and oxidative stress (Watson et al., 1998). Similarly, in *Listeria monocytogenes* cellular

morphology can change from bacilli to cocci during the transition to LTS phase, leading to an increased baro- and thermotolerance (Wen et al., 2011). Morphological changes were also observed during acid adaptation in the food grade bacterium *Propionibacterium freudenreichii*, with a shift from pleomorphic rods arranged in Chinese characters to bacteria with a segmented aspect and shorter cells (Jan et al., 2001).

Alterations of phenotype at the cell level can also have an impact at the colony level, namely in the size and color of the colonies (Helmus et al. 2012). A change in the size of colonies and a decline of the growth rate can be observed in *Bacillus subtilis*, during nutrient shortage (Chien; Hill; Levin, 2012). In *Streptococcus pyogenes*, after a long period (4 to 14 weeks) in stationary phase and nutrient shortage, culturable cells formed two colony phenotypes: atypical large colonies and microcolonies, which showed extensive changes at the proteome level compared to the parental strain (Wood et al., 2005).

**Table 2: Morphological alterations in bacterial cells or colonies.**

<b>Bacteria</b>	<b>Morphological alterations</b>	<b>Reference</b>
<i>Listeria monocytogenes</i>	The structure and the form of the cells change during the period of transition from early stationary phase to long-term survival	(Wen <i>et al.</i> 2011)
<i>Mycobacterium bovis</i>	In stressful conditions, <i>Mycobacterium bovis</i> displays different morphologies, such as L- forms (cell wall-deficient cells), as a way to survive in unfavorable environments	(Slavchev, Michailova and Markova 2013)
<i>Staphylococcus aureus</i>	Changes in morphology with a decrease in cell size between the hunger and survival phases	(Watson, Clements and Foster 1998)
<i>Streptococcus pyogenes</i>	After four weeks in stressful conditions, two different colony phenotypes are observed, most likely due to some selective pressure of the environment in two phases of life cycle: atypical large colonies and small colonies	(Wood 1981)



### **Growth rate changes**

Bacteria were early shown to be able to sense and recognize unfavorable conditions in the environment and thus to enter into a state where bacterial replication is inhibited until favorable conditions are re-established (in Wayne and Sohaskey, 2001: HJ Corper, Cohn ML, 1933). This capacity is an important strategy that allows bacterial survival. In *Bacillus subtilis*, alterations in cell division are directly correlated with nutrient availability and growth rate (Chien, Hill and Levin, 2012). To store energy, which correspond to a reduction in some metabolic pathways, such as the glycolytic pathway, in turn of decreasing the production of ATP (Savic and McShan 2012). As we can see other species, after a stress period, still require a period of latency for the morphological changes to occur, even after reestablishing favorable conditions (Roszak and Colwell 1987). This is because a part of the bacterial population needs more time to recover and return to being competitive after being exposed to the period of stress. The study of Geisel and colleagues shows the relationship between stress resistance and growth rate of various microorganisms. This study revealed that the cells remain in a vegetative state or enter into the dormant state, to decrease bacterial metabolism under unfavorable conditions, which allows these bacteria to adapt and survive in the environment (Geisel, Vilar and Rubi 2011).

### **Genetic adaptation**

In some bacterial species, LTS phase induces a growth advantage in stationary phase (GASP) phenomenon. GASP is defined as fluctuations in viable cell numbers during stationary phase growth, with a rapid loss of viability followed by an increase in viable counts (see above, Figure 3). In the stationary phase, the bacterial cells that undergo the GASP phenomenon are able to grow displace, the parent cells and become the dominant population in the culture. It is triggered by intense competition for nutrients and actually reflects the prevalent state of bacteria in their natural environment. GASP phenomenon is due to a series of genetic changes, occurring while the bacterial population enters the stationary phase. In parallel, physiological and structural alterations occur and induce a higher tolerance to diverse environmental stresses (Sewell et al., 2011). It was first reported and is well-documented in *E. coli*, the Gram negative model bacterium (Zinser, Erik R, Kolter, 2004; Zinster and Kolter, 1999; Bacun-Druzina et al. 2011). The GASP cells are able to transfer their advantage to

daughter cells, providing a competitive advantage (Finkel 2006a), in low concentrations of nutrients or in physical and chemical conditions that are unfavorable for growth. In contrast, only a few studies report on GASP phenomenon in Gram positive bacteria. One of the GASP associated phenotypes, namely the drop in cell counts at the beginning of stationary phases followed by an increase in cell counts, was shown in *Enterococcus faecium*. GASP however seems to be species-specific since it was not observed in *S. aureus* cultures (Sewell et al., 2011). In *Bacillus subtilis*, a phenomenon comparable to GASP was observed and linked to a hypermutable subpopulation that emerge during the stationary phase. This phenomenon, also called stationary-phase mutagenesis involves *comA* and *comK* genes (associated to prokaryotic differentiation) and is *recA*-independent. It leads to mutations that arise during the stationary-phase, in bacterial cells that do not divide and that are subjected to non lethal stresses, such as nutrient shortage (Sung and Yasbin, 2002).

### **Regulation of bacterial metabolism**

The maintenance and regulation of bacterial metabolism requires the action of several proteins involved in functions, such as transport and synthesis. In this section, we will highlight the factors that allow the control of these activities for the use of alternative energy sources, and the storage of nutrients and for the maintenance of intracellular homeostasis.

### **Use of atypical nitrogen and carbon source**

When the entry in LTS is due to nutritional shortage, some species adapt their metabolism to use other carbon and nitrogen sources. Such adaptation can be operated through the use of acetate or citrate as an alternative electron donor to produce energy, as shown in the Gram negative bacterium *Geobacter sulfurreducens* (Helmus et al. 2012). In Gram positive bacteria, *M. luteus* can use succinic acid as a substrate to produce energy. *Lactobacillus sakei* survival is linked to the degradation of arginine (Champomier Vergès et al. 1999), while *L. monocytogenes* can consume products from cell lysis to remain viable (Wen et al. 2013).

### **Nutrient Storage**

Some species of Gram-positive bacteria store various compounds, such as a polysaccharide, poly- $\gamma$ -hydroxybutyrate, lipids, polyphosphate or other elements, which

can be used to maintain their viability during prolonged periods of nutrient shortage (Wilkinson 1959).

In general, when bacterial cells are subjected to nutrient shortage in carbon, nitrogen and phosphorus, "survival" responses are induced. In such conditions, bacteria can use cell reserves like glycogen or polyphosphate (Carvalho et al. 2014) and/or cell lysis products in the LTS population (Roszak and Colwell 1987). Inorganic polyphosphate (polyP) is critical to bacterial survival in adverse conditions (Rao, Gómez-García and Kornberg 2009) as shown in *Mycobacterium sp.* During phosphate shortage, *Mycobacterium tuberculosis* induces the *ppk1* gene, which encodes a polyphosphate kinase and *relA*, which boosts the response to these stress so that the bacterial cells remain viable (Rifat, Bishai and Karakousis 2009).

In environments that do not present a broad range of available nutrients, bacteria are able to store some elements to prepare their survival phase, as shown in *Vibrio cholera* and *Mycobacterium tuberculosis*, which can store glucose in the form of glycogen (Bourassa and Camilli 2009; Amit Singh 2014). They can also do so when they are subjected to various stresses (nitrogen intermediates, reactive oxygen, low pH, and vital element starvation). When the carbon substrate is exhausted in the medium, these reserve substrates are utilized as energy source (Amit Singh 2014).

In LTS transition, *L. monocytogenes* upregulates the expression of genes related to energy metabolism, cell envelope structure, and transport. The upregulation of compatible solute transporters is thought to be involved in the accumulation of cellular solutes, which in turn lowers intracellular water activity and thus increases stress resistance (Greenblatt et al. 2004b; Wen et al. 2011).

### **Maintenance of pH and osmoprotection**

Bacterial adaptation to nutrient shortage results in the increase of some substrates, which can also be harmful to the cell. An increase in intracellular glucose can decrease bacterial survival as this will directly lower the pH of the medium because the bacteria cannot metabolize the large amounts of lactic acid produced. Some *S. pyogenes* strains are able to survive for long periods most likely thanks to mechanisms that allow them to escape these unfavorable conditions and remain viable, either by buffering systems or a similar mechanism for the maintenance of pH level (Savic and McShan 2012).

In *L. monocytogenes*, storage of osmoprotectants like glycine, betaine and trehalose allows viability. Also, upregulation of the transporter genes and the expression of genes involved in energy metabolic pathways is crucial for survival (Greenblatt et al. 2004a; Wen et al. 2011).

### **Different strategies for survival**

In this section, we discuss the best known mechanisms involved in LTS in Gram-positive bacteria, such as programmed cell death (PCD) and dormancy, which may help us understand this puzzle.

#### **Programmed cell death**

In Gram-positive bacteria, concepts such as programmed cell death have begun to emerge as a mechanism that allow bacterial populations to survive. Some authors suggest this as an evolutionary process that has allowed bacteria to adapt. Dynamic cellular stress responses cause destabilization bacterial population in varying conditions by requiring activation of quorum-sensing systems to trigger programmed cell death (PCD) (Finkel 2006c).

Quorum-sensing is observed in Gram-positive bacteria as a facilitator in response to stress conditions. The perception of stimuli in the medium activates specific signaling pathways, which initiates the process (Bharati et al. 2012). A large number of quiescent cells are responsible for the storage of carbon. In *L. monocytogenes*, this mechanism of quorum-sensing has been proposed as a trigger for activities that balance growth and death in the population by regulating gene expression between periods of long-term survival. This is possible because these bacteria are able to perceive the conditions to initiate programmed cell death, increasing the availability of nutrients to the surviving cells (Wen et al. 2013).

In addition to this mechanism, in this review, we will also focus on another very important phenomenon: dormancy.

#### **Dormancy**

Bacteria can to regulate their metabolic activity in order to maintain the viability of the population in periods without or without low cell multiplication and in periods of competition in extreme environments with low availability of nutrients (Ratcliff et al.

2013). This step of dormancy or quiescence corresponds to a “resting period” or a “period of reversible phenotypic change” and is triggered in response to extreme conditions of growth (Roszak and Colwell, 1987). It has been well-studied in pathogenic bacteria where, during an infection, the resulting slowly growing bacteria are able to tolerate both the host immune response and prolonged antibiotic exposure. This is typically the case in chronic infections such as tuberculosis caused by the Gram positive pathogen *M. tuberculosis*.

Dormancy (or quiescence) may be defined as a state of resistance to non-favorable conditions, or stresses where the viable population size is maintained throughout the period of stress, and where bacteria maintain a metabolic capacity, their membrane potential, and do not undergo obvious morphological differentiation (Rittershaus, Baek and Sassetti 2013a). During the dormancy period, cell replication and DNA synthesis are blocked, while the synthesis of RNA and proteins continues, so that bacterial metabolic activity remains viable until the stress is over (Roszak and Colwell, 1987). Besides *M. tuberculosis*, *Micrococcus luteus* was shown to enter in dormancy over long periods when subjected to starvation (Greenblatt et al. 2004b).

### **Viable but nonculturable state**

The entry in a Viable but nonculturable (VBNC) state is a strategy in response to adverse conditions described in various bacteria in situations of nutrient stress or atypical temperatures (Mulyukin et al. 2003). This concept emerged a few decades ago as an explanation for how bacteria can survive for long periods in adverse conditions (Roszak and Colwell, 1987). The VBNC state can be either a state in which bacteria await death or where they wait for a stimulus so that cellular metabolic activities can return to normal (Mulyukin et al. 2003).

It was found that in many species of pathogenic bacteria of VBNC state is present, so that they are able to express their virulence and remain viable. Furthermore, it was shown that bacteria have been VBNC the resuscitation promoting factor (Ramamurthy et al 2014).

### **Resuscitation**

Different methodologies that have been carried out in order to observe and characterize the phenomenon of “resuscitation bacterial”, seeks to restore conditions to

stabilize the population using suitable media, pH and temperature (Roszak and Colwell, 1987).

To get out of the state of dormancy, bacteria can initiate a form of recovery after a period of latency as reported in *M. luteus* (Greenblat et al. 2004b). This phase is named resuscitation and is also observed in *S. aureus*, which however requires a large amount and variety of nutrients to resume its metabolic activities (Diaper and Edwards 1994).

Resuscitation Promoting Factor (Rpf) is present in most bacteria with high G + C content. The Rpf are a group of secreted proteins, that belong to 29 families in genomes of actinomycetes (Dewi Puspita et al. 2013). In *M. tuberculosis*, it has been reported that anti-rpf antibodies are able to inhibit bacterial growth. Similarly, exogenous rpf added to old cultures leads to a reactivation of bacterial growth. Other models of infection *in vitro* were made and guaranteed performance of this group of proteins for resuscitation of cells (Kaprelyants et al. 2012).

Studies show that the analysis of the transcriptome of "resuscitated" cells of *L. monocytogenes* suggests an alteration in the metabolism to bacteria in their exponential phase, clearly showing that the bacterium is able to go through different conditions and remain viable (Wen et al. 2011). For this, it is necessary that genes associated with the synthesis of cell envelope components, proteins, amino acids and ribonucleotides are rapidly expressed (Wen et al. 2011).

### **What is to come? (Future directions)**

To date, most studies on long-term survival were carried out using classical microbiology techniques and mostly described the consequences of their phenotypes (size, color, persistence). However, is it possible, at a genomic level, to predict the ability of a bacterium to survive based on its gene content or based on the bacterial metabolic pathways?

New sequencing technologies and the advent of omic technologies, such as genomics, transcriptomics, proteomics and metagenomics, open avenues for research on the survival and adaptability of Gram-positive bacteria. Integrating information from different fields will allow to better understand how these mechanisms are regulated. This is of major interest to develop new products such as drugs, vaccines, or select more robust probiotics.

New research in bacterial survival is of great interest e.g. for the design of therapies that interfere with pathogenesis and persistence of bacteria, which are major concerns in clinical microbiology and food safety. A better understanding of LTS mechanisms could lead to the development of drugs that act even on dormant bacterial cells and impair the establishment of LTS in a pathological context, or in indwelling devices in hospitalized patients. Another application is to enhance LTS to optimize the efficiency of probiotic bacteria since some probiotic species are reportedly fragile (e.g. *Bifidobacterium* sp.) although their application requires a good survival during manufacturing and storage of products until the end consumer.

### **Author and Contributors**

The authors declare that there are no conflicts of interest.

### **Funding**

CAPES/COFECUB program supported this work. This work processes BEX 11517-12-3 and the Brazilian funding agencies CAPES and CNPq.

### **References**

- Amit Singh, Anil Kumar Gupta. Glycogenomics of *Mycobacterium tuberculosis*. *Mycobacterial Diseases*, v. 04, n. 06, 2014. Disponível em: <<http://www.omicsonline.org/open-access/glycogenomics-of-mycobacterium-tuberculosis-2161-1068-4-175.php?aid=35840>>. Acesso em: 25 set. 2015.
- Bhrati, B. K. et al. A full-length bifunctional protein involved in c-di-GMP turnover is required for long-term survival under nutrient starvation in *Mycobacterium smegmatis*. *Microbiology*, v. 158, n. Pt\_6, p. 1415–1427, 1 jun. 2012. Acesso em: 25 jun. 2014.
- Bourassa, Lori; Camilli, Andrew. Glycogen contributes to the environmental persistence and transmission of *Vibrio cholerae*. *Molecular Microbiology*, PMID: 19226328/PMCID: PMC2704980, v. 72, n. 1, p. 124–138, abr. 2009.
- Buchanan, R. E. Life Phases in a Bacterial Culture. *The Journal of Infectious Diseases*, v. 23, n. 2, p. 109–125, 1 ago. 1918.
- Carvalho, Mónica et al. Survival strategies of polyphosphate accumulating organisms and glycogen accumulating organisms under conditions of low organic loading. *Bioresource Technology*, v. 172, p. 290–296, nov. 2014.
- Champomier Vergès, Marie-Christine et al. Relationships between arginine degradation, pH and survival in *Lactobacillus sakei*. *FEMS Microbiology Letters*, v. 180, n. 2, p. 297–304, nov. 1999. Acesso em: 25 jun. 2014.

- Chaussabel, Damien; Sher, Alan. Mining microarray expression data by literature profiling. *Genome Biology*, PMID: 12372143PMCID: PMC134484, v. 3, n. 10, p. research0055.1–research0055.16, 2002.
- Chien, An-Chun; Hill, Norbert S.; Levin, Petra Anne. Cell Size Control in Bacteria. *Current Biology*, v. 22, n. 9, p. R340–R349, 8 maio 2012.
- Dewi Puspita, Indun et al. Resuscitation Promoting Factor (Rpf) from *Tomitella biformata* AHU 1821T Promotes Growth and Resuscitates Non-Dividing Cells. *Microbes and Environments*, v. 28, n. 1, p. 58–64, 2013. Acesso em: 12 set. 2014.
- Diaper, J. P.; Edwards, C. Survival of *Staphylococcus aureus* in lakewater monitored by flow cytometry. *Microbiology (Reading, England)*, PMID: 7512871, v. 140 ( Pt 1), p. 35–42, jan. 1994.
- Finkel, Steven E. Long-term survival during stationary phase: evolution and the GASP phenotype. *Nature Reviews Microbiology*, v. 4, n. 2, p. 113–120, fev. 2006a. Acesso em: 16 abr. 2014.
- Finkel, Steven E. Long-term survival during stationary phase: evolution and the GASP phenotype. *Nature Reviews Microbiology*, v. 4, n. 2, p. 113–120, fev. 2006b. Acesso em: 25 jun. 2014.
- Geisel, Nico; Vilar, Jose M. G.; Rubi, J. Miguel. Optimal Resting-Growth Strategies of Microbial Populations in Fluctuating Environments. *PLoS ONE*, v. 6, n. 4, p. e18622, 15 abr. 2011.
- Greenblatt, C. L. et al. *Micrococcus luteus* - Survival in Amber. *Microbial Ecology*, v. 48, n. 1, p. 120–127, 28 maio 2004a.
- Greenblatt, C.L. et al. *Micrococcus luteus* - Survival in Amber. *Microbial Ecology*, v. 48, n. 1, p. 120–127, set. 2004b. Acesso em: 28 abr. 2014.
- Helmus, Ruth A. et al. Growth advantage in stationary-phase (GASP) phenotype in long-term survival strains of *Geobacter sulfurreducens*. *FEMS Microbiology Ecology*, v. 79, n. 1, p. 218–228, jan. 2012. Acesso em: 16 abr. 2014.
- Jan, Gwénaél et al. Changes in Protein Synthesis and Morphology during Acid Adaptation of *Propionibacterium freudenreichii*. *Applied and Environmental Microbiology*, PMID: 11319077PMCID: PMC92832, v. 67, n. 5, p. 2029–2036, maio 2001.
- Markova, Nadya; Slavchev, Georgi; Michailova, Lilia. Unique biological properties of *Mycobacterium tuberculosis* L-form variants: impact for survival under stress. *International Microbiology: The Official Journal of the Spanish Society for Microbiology*, PMID: 22847267, v. 15, n. 2, p. 61–68, jun. 2012.
- Mulyukin, Andrei L. et al. Formation of resting cells by non-spore-forming microorganisms as a strategy of long-term survival in the environment. 30 jan. 2003, [S.l.: s.n.], 30 jan. 2003. p. 208–218. Disponível em: <<http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=889034>>. Acesso em: 25 jun. 2014.



Palop, Alfredo; Mañas, Pilar; Condón, Santiago. Sporulation temperature and heat resistance of *Bacillus* Spores: a Review. *Journal of Food Safety*, v. 19, n. 1, p. 57–72, abr. 1999.

Ramamurthy, Thandavarayan et al. Current Perspectives on Viable but Non-Culturable (VBNC) Pathogenic Bacteria. *Frontiers in Public Health*, v. 2, 31 jul. 2014. Disponível em: <<http://journal.frontiersin.org/article/10.3389/fpubh.2014.00103/abstract>>. Acesso em: 7 abr. 2015.

Rao, Narayana N.; Gómez-García, María R.; Kornberg, Arthur. Inorganic Polyphosphate: Essential for Growth and Survival. *Annual Review of Biochemistry*, v. 78, n. 1, p. 605–647, 2009.

Ratcliff, William C. et al. Disentangling direct and indirect fitness effects of microbial dormancy. *The American Naturalist*, v. 182, n. 2, p. 147–156, ago. 2013. Acesso em: 25 jun. 2014.

Rifat, Dalin; Bishai, William R.; Karakousis, Petros C. Phosphate depletion: a novel trigger for *Mycobacterium tuberculosis* persistence. *The Journal of Infectious Diseases*, PMID: 19686042, v. 200, n. 7, p. 1126–1135, 1 out. 2009.

Rittershaus, Emily S.C.; Baek, Seung-Hun; Sasseti, Christopher M. The Normalcy of Dormancy: Common Themes in Microbial Quiescence. *Cell Host & Microbe*, v. 13, n. 6, p. 643–651, jun. 2013. Acesso em: 13 ago. 2014.

Rozzak, D B; Colwell, R R. Survival strategies of bacteria in the natural environment. *Microbiological Reviews*, PMID: 3312987PMCID: PMC373117, v. 51, n. 3, p. 365–379, set. 1987.

Savic, D. J.; Mcshan, W. M. Long-term survival of *Streptococcus pyogenes* in rich media is pH-dependent. *Microbiology*, v. 158, n. Pt\_6, p. 1428–1436, 1 jun. 2012. Acesso em: 25 jun. 2014.

Sewell, D.; Laird, K.; Phillips, C. Growth advantage in stationary phase phenomenon in Gram-positive bacteria. *Journal of Hospital Infection*, v. 78, n. 1, p. 73–75, maio 2011.

S. Kaprelyants, Arseny et al. Resuscitation-promoting Factors (Rpf): In Search of Inhibitors. *Protein & Peptide Letters*, v. 19, n. 10, p. 1026–1034, 1 set. 2012. Acesso em: 12 set. 2014.

Slavchev, Georgi; Michailova, Lilia; Markova, Nadya. Stress-induced L-forms of *Mycobacterium bovis*: a challenge to survivability. *The New Microbiologica*, PMID: 23686122, v. 36, n. 2, p. 157–166, abr. 2013.

Sung, Huang-Mo; Yasbin A, Ronald E. Adaptive, or stationary-phase, mutagenesis, a component of bacterial differentiation in *Bacillus subtilis*. *Journal of Bacteriology*, PMID: 12270822PMCID: PMC139596, v. 184, n. 20, p. 5641–5653, out. 2002.

Watson, S. P.; Clements, M. O.; Foster, S. J. Characterization of the starvation-survival response of *Staphylococcus aureus*. *Journal of Bacteriology*, PMID: 9537371 PMCID: PMC107086, v. 180, n. 7, p. 1750–1758, abr. 1998.

Wayne, L. G.; Sohaskey, C. D. Nonreplicating persistence of mycobacterium tuberculosis. *Annual Review of Microbiology*, PMID: 11544352, v. 55, p. 139–163, 2001.

Wen, J. et al. Transcriptomic Response of *Listeria monocytogenes* during the Transition to the Long-Term-Survival Phase. *Applied and Environmental Microbiology*, v. 77, n. 17, p. 5966–5972, 1 set. 2011. Acesso em: 25 jun. 2014.

Wen, Jia et al. *Listeria monocytogenes* responds to cell density as it transitions to the long-term-survival phase. *International Journal of Food Microbiology*, v. 165, n. 3, p. 326–331, ago. 2013. Acesso em: 25 jun. 2014.

Wilkinson, J. F. The problem of energy-storage compounds in bacteria. *Experimental Cell Research, The Cytochemistry of Enzymes and Antigens*. v. 7, Supplement, p. 111–130, 1959.

Wood, D. N. et al. Persistence of *Streptococcus pyogenes* in Stationary-Phase Cultures. *Journal of Bacteriology*, v. 187, n. 10, p. 3319–3328, 15 maio 2005. Acesso em: 25 abr. 2014.

Zinser, Erik R; Kolter, Roberto. *Escherichia coli* evolution during stationary phase. *Research in Microbiology*, v. 155, n. 5, p. 328–336, jun. 2004. Acesso em: 25 jun. 2014.

Zinser, Erik R; Kolter, Roberto. Mutations enhancing amino acid catabolism confer a growth advantage in stationary phase. *Journal of Bacteriology*, PMID: 10482523/PMCID: PMC94102, v. 181, n. 18, p. 5800–5807, set. 1999.

## Anexo 7: Curriculum vitae

## Annexe 7 : Curriculum vitae

Address to access this CV: <http://lattes.cnpq.br/2016464855117057>

### *Personal Information*

**Full name** Flavia Figueira Aburjaile

**Name used in Bibliographic Citations** ABURJAILE, F. F.; FIGUEIRA, F.A.; ABURJAILE, F.; FIGUEIRA, F.; FLÁVIA; FIGUEIRA ABURJAILE, F.; Flávia Aburjaile; ABURJAILE, FF; Aburjaile, Flávia; ABURJAILE, FLÁVIA FIGUEIRA

**Professional Address** Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas  
Av. Antônio Carlos, 6627, Laboratório de Genética Celular e Molecular,  
Bloco Q3, sala 259  
Pampulha - Belo Horizonte  
31270910, MG - Brazil  
Phone number: 31 3409-2873

**e-mail** faburjaile@gmail.com

### *Formal Education*

**2011 - 2015** Doctorate in Bioinformatics.  
Universidade Federal de Minas Gerais, UFMG, Belo Horizonte, Brazil  
(**Co-guardianship**) Institut National de la Recherche Agronomique  
(Advisor: Yves Le Loir)  
Title: Mecanismos moleculares de sobrevivência em longo prazo em *Propionibacterium freudenreichii*  
Advisor: Vasco Ariston de Carvalho Azevedo  
Scholarship from: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior/COFECUB  
Knowledge areas: Bioinformatics and Molecular Genetics in Microorganisms.

**2007 - 2011** Bachelor's in Biomedicine.  
Universidade Fundação Mineira de Educação e Cultura, FUMEC, Belo Horizonte, Brasil  
Study abroad sandwich option: Université de Provence Aix Marseille I  
(Advisor : Jean-Pierre Roll)  
Title: Teste do efeito dos probióticos frente à *Pseudomonas aeruginosa*  
Advisor: Adriana dos Santos e Flávio Henrique Barbosa

### ***Complementary Education***

- 2014 - 2014** Short Course in Advanced seq annotation and use of pipelines EGene2.  
Universidade Federal de Minas Gerais, UFMG, Belo Horizonte, Brasil
- 2014 - 2014** Short Course in PATRIC.  
Universidade Federal de Minas Gerais, UFMG, Belo Horizonte, Brasil
- 2014 - 2014** Short Course in Genomics for Working with Pathogens.  
EMBL European Bioinformatics Institute, EMBL-EBI, UFMG, Belo Horizonte, Brasil
- 2014 - 2014** Short Course in GO terms and annotations.  
Universidade Federal de Minas Gerais, UFMG, Belo Horizonte, Brasil
- 2012 - 2012** Short Course in Introduction to BioNumerics program.  
Universidade Federal de Minas Gerais, UFMG, Belo Horizonte, Brasil
- 2012 - 2012** Short Course in Introduction to Perl languages.  
Universidade Federal de Minas Gerais, UFMG, Belo Horizonte, Brasil
- 2012 - 2012** Short Course in Molecular Methods for Microbial Diagnosis, Genomic.  
Universidade Federal de Minas Gerais, UFMG, Belo Horizonte, Brasil
- 2012 - 2012** Short Course em Rna-seq course.  
Universidade Federal de Minas Gerais, UFMG, Belo Horizonte, Brasil
- 2012 - 2012** Short Course em International Course in Automated Functional Annot.  
Centro de Investigaciones Príncipe Felipe, CIPF, Espanha
- 2011 - 2011** Short Course in Next Generation Sequencing: SOLiD.  
Universidade Federal do Paraná, UFPR, Curitiba, Brasil
- 2011 - 2011** Short Course in assembly and annotation of bacterial genomes.  
Universidade Federal de Minas Gerais, UFMG, Belo Horizonte, Brasil
- 2011 - 2011** Short Course in II Winter Course of Biochemistry and Molecular Biology.  
Universidade Federal do Paraná, UFPR, Curitiba, Brasil
- 2011 - 2011** Short Course in molecular biology transformation.  
Universidade Federal do Paraná, UFPR, Curitiba, Brasil
- 2011 - 2011** Short Course in IV Winter Course of Genetics.  
Universidade Federal do Paraná, UFPR, Curitiba, Brasil
- 2010 - 2010** Short Course in Stage intensif de Langue française.

Université de Provence Aix Marseille I, Aix-Marceille I, Marseille ,  
França

**2009 - 2009** Short Course in Arterial Blood Gas Analysis.  
Medical Association of Minas Gerais, AMMG, Brazil, AMMG, Brasil

**2008 - 2009** French. Alliance Française, BH, Brazil.

### *Areas of expertise*

1. Bioinformatics
2. Microbiology
3. Genetics
4. Parasitology

### *Languages*

1. English Understanding Functional, Speaking Functional, Writing Basic, Reading  
Fluent
2. Spanish Understanding Functional, Speaking Functional, Writing Basic, Reading  
Fluent
3. French Understanding Fluent, Speaking Functional, Writing Functional, Reading  
Fluent

### **Prêmios e títulos**

**2013** BioVis 2013 - Data Contest Biology Experts Pick, IEEE Computer Society

### *S,T &A Production*

#### **Bibliographic Production**

#### **Articles Published in Scientific Journals**

1. BARBOSA, EUDES G. V., **ABURJAILE, F. F.**, RAMOS, ROMMEL T.J., CARNEIRO, ADRIANA R., LOIR, Y. L., BAUMBACH, J., MIYOSHI, ANDERSON, SILVA, ARTUR, AZEVEDO, VASCO. Value of a newly sequenced bacterial genome. World Journal of Biological Chemistry. , v.5, p.161 - 168, 2014.

2. SOARES, SIOMAR C., SILVA, ARTUR, TROST, EVA, BLOM, JOCHEN, RAMOS, ROMMEL, CARNEIRO, ADRIANA, ALI, AMJAD, SANTOS, ANDERSON R., PINTO, ANNE C., DINIZ, CARLOS, BARBOSA, EUDES G. V., DORELLA, FERNANDA A., **ABURJAILE, FLÁVIA**, ROCHA, FLÁVIA S., NASCIMENTO, KARINA K. F., GUIMARÃES, LUÍS C., ALMEIDA, SINTIA, HASSAN, SYED S., BAKHTIAR, SYEDA M., PEREIRA, ULISSES P., ABREU, VINICIUS A. C., SCHNEIDER, MARIA P. C., MIYOSHI, ANDERSON, TAUCH, ANDREAS,

AZEVEDO, VASCO. The Pan-Genome of the Animal Pathogen *Corynebacterium pseudotuberculosis* Reveals Differences in Genome Plasticity between the Biovar ovis and equi Strains. *Plos One.* , v.8, p.e53818-e53818, 2013.

3. DORELLA, FERNANDA A, GALA-GARCIA, ALFONSO, PINTO, ANNE C, SARROUH, BOUTROS, ANTUNES, CAMILA A, RIBEIRO, DAYANA, **Aburjaile, Flávia**, FIAUX, KARINA K, Guimarães, Luis C, SEYFFERT, NUBIA, EL-AOUAR, RACHID A, SILVA, RENATA, HASSAN, SYED S, CASTO, THIAGO LP, MARQUES, WANDERSON S, RAMOS, ROMMEL, CARNEIRO, ADRIANA, SÁ, PABLO DE, MIYOSHI, ANDERSON, AZEVEDO, VASCO, SILVA, ARTUR. Progression of 'OMICS' methodologies for understanding the pathogenicity of *Corynebacterium pseudotuberculosis*: the Brazilian experience. *Computational and Structural Biotechnology Journal.* , v.6, p.1 - 7, 2013.

4. RAMOS, ROMMEL THIAGO JUCÁ, CARNEIRO, ADRIANA RIBEIRO, SOARES, SIOMAR DE CASTRO, SANTOS, ANDERSON RODRIGUES DOS, ALMEIDA, SINTIA, GUIMARÃES, LUIS, **FIGUEIRA, FLÁVIA**, BARBOSA, EUDES, TAUCH, ANDREAS, AZEVEDO, VASCO, SILVA, ARTUR. Tips and tricks for the assembly of a *Corynebacterium pseudotuberculosis* genome using a semiconductor sequencer. *Microbial Biotechnology (Online).* , v.n/a, p.n/a - n/a, 2012.

5. Soares, Siomar C., Trost, Eva, Ramos, Rommel T.J., Carneiro, Adriana R., Santos, Anderson R., Santos, A.R., Pinto, Anne C., **ABURJAILE, F. F.**, Barbosa, Eudes, Ali, Amjad, DINIZ, CARLOS A.A., Hassan, Syed S., FIAUX, KARINA, Guimarães, Luis C., Bakhtiar, Syeda M., PEREIRA, ULISSES, Almeida, Sintia S., Abreu, Vinícius A.C., ROCHA, FLÁVIA S., DORELLA, FERNANDA A., Miyoshi, Anderson, Silva, Artur, Azevedo, Vasco, Tauch, Andreas. Genome sequence of *Corynebacterium pseudotuberculosis* biovar equi strain 258 and prediction of antigenic targets to improve biotechnological vaccine production. *Journal of Biotechnology.* , v.166, p.1 - 7, 2012.

6. SILVA, A., LOPES, T., RAMOS, R. T. J, CARNEIRO, A., DORELLA, F., Rocha, F. S., SANTOS, A. R., LIMA, A. R. J., GUIMARAES, L. C., Barbosa, E. G. V., RIBEIRO, D., FIAUX, K., DINIZ, C. A. A., ABREU, V. A. C., ALMEIDA, S. S., HASSAN, S., ALI, A., BAKHTIAR, S., **ABURJAILE, F. F.**, PINTO, A. C., Soares, S. C., PEREIRA, U., Schneider, M. P. C, MIYOSHI, A., EDMAN, J., SPIER, S, AZEVEDO, V.. Complete genome sequence of *Corynebacterium pseudotuberculosis* Cp267 strain isolated from a Llama. *Journal of Bacteriology (Print).* , v.194, p.3567 - 3568, 2012.

7. Lopes, T., Silva, A., Thiago, R., Carneiro, A., Dorella, F. A., Rocha, F. S., dos Santos, A. R., Lima, A. R. J., Guimaraes, L. C., Barbosa, E. G. V., **BARBOSA, E. G. V.**, Ribeiro, D., Fiaux, K. K., Diniz, C. A. A., de Abreu, V. A. C., de Almeida, S. S., Hassan, S. S., Ali, A., **ABURJAILE, F. F.**, Bakhtiar, S. M., Pinto, A. C., Soares, S. d. C., Pereira, U. d. P., Schneider, M. P. C., Miyoshi, A., Edman, J., Spier, S., Azevedo, V. Complete Genome Sequence of *Corynebacterium pseudotuberculosis* Strain Cp267, Isolated from a Llama. *Journal of Bacteriology (Print).* , v.194, p.3567 - 3568, 2012.

8. Pethick, F. E., Lainson, A. F., YAGA, R., FLOCKHART, A., Smith, D. G. E., DONACHIE, W., Cerdeira, L. T., Silva, A., BOL, E., LOPES, T. S., Barbosa, M. S., PINTO, A. C., dos Santos, A. R., SOARES, S. C., Almeida, S. S., Guimaraes, L. C.,

**ABURJAILE, F. F.**, Abreu, V. A. C., RIBEIRO, D., FIAUX, K. K., Diniz, C. A. A., BARBOSA, E. G. V., Pereira, U. P., Hassan, S. S., Ali, A., Bakhtiar, S. M., DORELLA, F. A., DORELLA, F. A., CARNEIRO, A. R., Ramos, R. T. J., Rocha, F. S., Schneider, M. P. C., MIYOSHI, A., AZEVEDO, V., Fontaine, M. C. Complete Genome Sequence of *Corynebacterium pseudotuberculosis* Strain 1/06-A, Isolated from a Horse in North America. *Journal of Bacteriology* (Print). , v.194, p.4476 - 4476, 2012.

9. Pethick, F. E., Lainson, A. F., Yaga, R., Flockhart, A., Smith, D. G. E., Donachie, W., CERDEIRA, L. T., SILVA, A., Bol, E., LOPES, T. S., Barbosa, M. S., PINTO, A. C., dos Santos, A. R., SOARES, S. C., ALMEIDA, S. S., Guimaraes, L. C., **ABURJAILE, F. F.**, ABREU, V. A. C., RIBEIRO, D., Fiaux, K. K., Diniz, C. A. A., BARBOSA, E. G. V., Pereira, U. P., Hassan, S. S., ALI, A., Bakhtiar, S. M., DORELLA, F. A., CARNEIRO, A. R., RAMOS, R. T. J., Rocha, F. S., SCHNEIDER, M. P. C., MIYOSHI, A., AZEVEDO, V., AZEVEDO, V., Fontaine, M. C. Complete Genome Sequences of *Corynebacterium pseudotuberculosis* Strains 3/99-5 and 42/02-A, Isolated from Sheep in Scotland and Australia, Respectively. *Journal of Bacteriology* (Print). , v.194, p.4736 - 4737, 2012.

10. Santos, A.R., Carneiro, A. R., Ramos, R. T. J., DALL'AGNOL, H., PINTO, A. C., de Castro Soares, S., SANTOS, A. R., Guimaraes, L. C., Almeida, S. S., BARAUNA, R. A., DAS GRACAS, D. A., FRANCO, L. C., Ali, A., Hassan, S. S., NUNES, C. I. P., Barbosa, M. S., **ABURJAILE, F. F.**, Fiaux, K. K., Barbosa, E. G. V., Bakhtiar, S. M., VILELA, D., NOBREGA, F., DOS SANTOS, A. L., CAREPO, M. S. P., Azevedo, V., Schneider, M. P. C., PELLIZARI, V. H., SILVA, A. Genome Sequence of *Exiguobacterium antarcticum* B7, Isolated from a Biofilm in Ginger Lake, King George Island, Antarctica. *Journal of Bacteriology* (Print). , v.194, p.6689 - 6690, 2012.

11. SANTOS, A. R., CARNEIRO, A. R., GALA-GARCIA, A., PINTO, A. C., BARH, D., BARBOSA, E. G. V., **ABURJAILE, F. F.**, DORELLA, F. A., ROCHA, F. S., GUIMARAES, L. C., Zurita-Turk, Meritxell, RAMOS, R. T. J., ALMEIDA, S. S., ALMEIDA, S., SOARES, S. C., PEREIRA, U. P., ABREU, V. A. C., SILVA, A., MIYOSHI, A., AZEVEDO, V.. The *Corynebacterium pseudotuberculosis* in silico predicted pan-exoproteome. *BMC Genomics*. , v.13, p.S5 - S6, 2012.

12. CERDEIRA, L., Schneider, M. P. C, PINTO, A. C., ALMEIDA, S. S., SANTOS, A. R., Barbosa, E. G. V., ALI, A., **ABURJAILE, F. F.**, ABREU, V. A. C., GUIMARAES, L. C., Soares, S. C., DORELLA, F., Rocha, F. S., BOL, E., LOPES, T., Barbosa, M. S., CARNEIRO, A., Juca Ramos, R. T., Coimbra, N. A. d. R, LIMA, A. R. J., Barh, D, Tiwari, S., RAJA, R., ZAMBARE, V., Ghosh, P., Trost, E., Tauch, A, MIYOSHI, A., AZEVEDO, V., SILVA, A.. Complete Genome Sequence of *Corynebacterium pseudotuberculosis* Strain CIP 52.97, Isolated from a Horse in Kenya. *Journal of Bacteriology* (Print). , v.193, p.7025 - 7026, 2011.

13. HASSAN, SYED SHAH, GUIMARÃES, LUIS CARLOS, PEREIRA, ULISSES DE PÁDUA, ISLAM, ARSHAD, ALI, AMJAD, BAKHTIAR, SYEDA MARRIAM, RIBEIRO, DAYANA, RODRIGUES DOS SANTOS, ANDERSON, SOARES, SIOMAR DE CASTRO, DORELLA, FERNANDA, PINTO, ANNE CYBELLE, SCHNEIDER, MARIA PAULA CRUZ, BARBOSA, MARIA SILVANIRA, ALMEIDA, SÍNTIA, ABREU, VINÍCIUS, **ABURJAILE, FLÁVIA**, CARNEIRO, ADRIANA RIBEIRO, CERDEIRA, LOUISE TEIXEIRA, FIAUX, KARINA,

BARBOSA, EUDES, DINIZ, CARLOS, ROCHA, FLAVIA S., RAMOS, ROMMEL THIAGO JUCÁ, JAIN, NEHA, TIWARI, SANDEEP, BARH, DEBMALYA, MIYOSHI, ANDERSON, MÜLLER, BORNA, SILVA, ARTUR, AZEVEDO, VASCO. Complete genome sequence of *Corynebacterium pseudotuberculosis* biovar ovis strain P54B96 isolated from antelope in South Africa obtained by rapid next generation sequencing technology. *STAND GENOMIC SCI.* , v.7, p.189 - 199, 2012.

14. SILVA, A., RAMOS, R. T. J., RIBEIRO CARNEIRO, A., CYBELLE PINTO, A., de Castro Soares, S., RODRIGUES SANTOS, A., SILVA ALMEIDA, S., Guimaraes, L. C., **FIGUEIRA ABURJAILE, F.**, VIEIRA BARBOSA, E. G., ALVES DORELLA, F., SOUZA ROCHA, F., SOUZA LOPES, T., KAWASAKI, R., GOMES SA, P., DA ROCHA COIMBRA, N. A., TEIXEIRA CERDEIRA, L., SILVANIRA BARBOSA, M., CRUZ SCHNEIDER, M. P., MIYOSHI, A., SELIM, S. A. K., MOAWAD, M. S., AZEVEDO, V. Complete Genome Sequence of *Corynebacterium pseudotuberculosis* Cp31, Isolated from an Egyptian Buffalo. *Journal of Bacteriology (Print).* , v.194, p.6663 - 6664, 2012.

15. Ramos, R. T. J., Silva, A., CARNEIRO, A. R., PINTO, A. C., Soares, S. d. C., Santos, A. R., Almeida, S. S., Guimaraes, L. C., **ABURJAILE, F. F.**, BARBOSA, E. G. V., DORELLA, F. A., DORELLA, F. A., Rocha, F. S., Cerdeira, L. T., Barbosa, M. S., Tauch, A., EDMAN, J., SPIER, S., MIYOSHI, A., Schneider, M. P. C., AZEVEDO, V. Genome Sequence of the *Corynebacterium pseudotuberculosis* Cp316 Strain, Isolated from the Abscess of a Californian Horse. *Journal of Bacteriology (Print).* , v.194, p.6620 - 6621, 2012.

16. Hassan, S. S., SCHNEIDER, M. P. C., RAMOS, R. T. J., CARNEIRO, A. R., RANIERI, A., Guimaraes, L. C., ALI, A., Bakhtiar, S. M., Pereira, U. d. P., dos Santos, A. R., Soares, S. d. C., DORELLA, F., PINTO, A. C., RIBEIRO, D., Barbosa, M. S., ALMEIDA, S., ABREU, V., **ABURJAILE, F.**, FIAUX, K., BARBOSA, E., DINIZ, C., ROCHA, F. S., SAXENA, R., Tiwari, S., ZAMBARE, V., Ghosh, P., PACHECO, L. G. C., DOWSON, C. G., KUMAR, A., BARH, D., MIYOSHI, A., AZEVEDO, V., SILVA, A.. Whole-Genome Sequence of *Corynebacterium pseudotuberculosis* Strain Cp162, Isolated from Camel. *Journal of Bacteriology (Print).* , v.194, p.5718 - 5719, 2012.

17. OLIVEIRA, L. C., SARAIVA, T. D. L., SOARES, S. C., RAMOS, R. T. J., SA, P. H. C. G., CARNEIRO, A. R., MIRANDA, F., FREIRE, M., RENAN, W., JUNIOR, A. F. O., SANTOS, A. R., PINTO, A. C., SOUZA, B. M., CASTRO, C. P., Diniz, C. A. A., ROCHA, C. S., MARIANO, D. C. B., DE AGUIAR, E. L., FOLADOR, E. L., Barbosa, E. G. V., **ABURJAILE, F. F.**, GONCALVES, L. A., GUIMARAES, L. C., AZEVEDO, M., AGRETI, P. C. M., SILVA, R. F., Tiwari, S., ALMEIDA, S. S., Hassan, S. S., PEREIRA, V. B., ABREU, V. A. C., Pereira, U. P., Dorella, F. A., CARVALHO, A. F., PEREIRA, F. L., LEAL, C. A. G., FIGUEIREDO, H. C. P., SILVA, A., MIYOSHI, A., AZEVEDO, V. Genome Sequence of *Lactococcus lactis* subsp. *lactis* NCDO 2118, a GABA-Producing Strain. *Genome Announcements.* , v.2, p.e00980-14 - e00980-14, 2014.

18. SANTOS, ANDERSON R, CARNEIRO, ADRIANA, GALA-GARCÍA, ALFONSO, PINTO, ANNE, BARH, DEBMALYA, BARBOSA, EUDES, **ABURJAILE, FLÁVIA**, DORELLA, FERNANDA, ROCHA, FLÁVIA, GUIMARÃES, LUIS, Zurita-Turk, Meritxell, RAMOS, ROMMEL, ALMEIDA,



SINTIA, SOARES, SIOMAR, PEREIRA, ULISSES, ABREU, VINÍCIUS C, SILVA, ARTUR, MIYOSHI, ANDERSON, AZEVEDO, VASCO. The *Corynebacterium pseudotuberculosis* in silico predicted pan-exoproteome. BMC Genomics. , v.13, p.S6, 2012.

19. ABURJAILE, S. B., **ABURJAILE, F. F.**, BARBOSA, F. H. F.;Barbosa, Flávio H. F.;BARBOSA, F. Variação da concentração de minerais em aquíferos confinados de indústria de exploração de águas minerais naturais no município de Aparecida-SP. Ciência Equatorial. , v.1, p.13 - 17, 2011.

20. BARBOSA, F. H. F.;Barbosa, Flávio H. F.;BARBOSA, F, BARBOSA, L., BAMBIRRA, Leandro Henrique Silva, **ABURJAILE, F. F.** Produção de substâncias envolvidas no fenômeno de antagonismo bacteriano. Revista de Biologia e Ciências da Terra. , v.11, p.1 - 10, 2011.

21. BARBOSA, F. H. F.;Barbosa, Flávio H. F.;BARBOSA, F, BARBOSA, L., BAMBIRRA, Leandro Henrique Silva, **ABURJAILE, F. F.** Probióticos microrganismos a favor da vida. Revista de Biologia e Ciências da Terra. , v.11, p.11 - 21, 2011.

22. PEREIRA, ULISSES DE PÁDUA, RODRIGUES DOS SANTOS, ANDERSON, HASSAN, SYED SHAH, **ABURJAILE, FLÁVIA FIGUEIRA**, SOARES, SIOMAR DE CASTRO, RAMOS, ROMMEL THIAGO JUCÁ, CARNEIRO, ADRIANA RIBEIRO, GUIMARÃES, LUÍS CARLOS, SILVA DE ALMEIDA, SINTIA, DINIZ, CARLOS AUGUSTO ALMEIDA, BARBOSA, MARIA SILVANIRA, GOMES DE SÁ, PABLO, ALI, AMJAD, BAKHTIAR, SYEDA MARRIAM, DORELLA, FERNANDA ALVES, ZERLOTINI, ADHEMAR, ARAÚJO, FLÁVIO MARCOS GOMES, LEITE, LAURA RABELO, OLIVEIRA, GUILHERME, MIYOSHI, ANDERSON, SILVA, ARTUR, AZEVEDO, VASCO, FIGUEIREDO, HENRIQUE CÉSAR PEREIRA. Complete genome sequence of *Streptococcus agalactiae* strain SA20-06, a fish pathogen associated to meningoencephalitis outbreaks. Standards in Genomic Sciences. , v.8, p.188 - 197, 2013.

#### Articles accepted for publication

1. **ABURJAILE, FF**, MADEC, M., PARAYRE, S., AZEVEDO, V., MIYOSHI, A., LE LOIR, Y. , FALENTIN, H.. *Propionibacterium freudenreichii* long-term survival in nutrient shortage. Journal of Applied Microbiology (Print). , 2015.

#### Book chapters published

1. Roukos, Dimitrios H, SILVA, ARTUR, Carneiro, Adriana R, **Aburjaile, Flávia**, Guimarães, Luis C, Ramos, Rommel TJ, Castro, Thiago LP, Abreu, Vinicius, Silva, Wanderson M, Schneider, Paula, AZEVEDO, VASCO. Comprehensive whole genome and transcriptome analysis for novel diagnostics In: Next-Generation Sequencing & Molecular Diagnostics. 1 ed.London: Future Medicine, p. 64-76, 2013

Home page: <http://www.futuremedicine.com/doi/book/10.2217/9781780841861>

2. **ABURJAILE, F. F.**, SANTANA, M. P., VIANA, M. V. C., Silva, Wanderson M, FOLADOR, E., SILVA, A., AZEVEDO, V.. Genomics In: A Textbook of Biotechnology. 1 ed. Texas : SMGROUP, 2015, p. 20-40.  
ISBN: 9780996274531

### ***Book chapters accepted for publication***

1. Santana, M., **Aburjaile, F.F.**, Parise, M., Tiwari, S., Pinto, A and Azevedo V. RNA-seq: Reveling biological insights in Bacteria, InTech - Next Generation Sequencing, 2015.

### **Articles in Newspapers/Magazines**

1. Soares, S. C., Silva, A, RAMOS, R. T. J, CERDEIRA, L., ALI, A., SANTOS, A. R., PINTO, A. C., Cassiano, AAM, FERNANDES, J. G. A., ABURJAILE, F. F., CARNEIRO, A. R., GUIMARAES, L. C., Barbosa, E. G. V., ALMEIDA, S. S., ABREU, V. A. C., MIYOSHI, A., AZEVEDO, V.. Plasticidade Genômica e Evolução Bacteriana. Microbiologia in Foco, 26 CBM. Foz do Iguaçu, p.31 - 38, 2011.

### **Innovation**

#### ***Research Project***

**2009 - 2009** Projeto ManAli (Food handlers)  
Concluded

Integrantes: Flavia Figueira Aburjaile, Fernanda Diniz Prates, Juliana Fernandes, Kenia Cassia Pinto Gazola, Adriana dos Santos

#### ***Event Production***

##### **1. ABURJAILE, F. F.**

Member of the Organizing Committee of the II Congress of Microbiology and Parasitology “Oswino Penna Sobrinho”, 2009.

##### **2. ABURJAILE, F. F.**

Member of the Organizing Committee of Academic Week the Health Sciences College in Universidade Fumec, 2008.

### ***Complementary education in France :***

#### **Poster presentation:**

- November 04 to 06 (2013) → 2<sup>nd</sup> Colloque de Génomique Environnementale (Rennes)

#### **Courses Training:**

- October 24 to 26 (2012) → Blast2Go (Automatic Functional Annotation and Data Mining) –Valence – Spain
- April 18 and 19 (2013) → Training: MISTER (Rennes)
- 21 and May 22 (2013) → Training: MIGALE - GALAXY (Jouy-en-Josas)

**French and English courses:**

- October 1<sup>st</sup> to December 21 (2012) → French course (B1) at CIREFE univ. Rennes 2 (4 hours per week)
- February 4 to May 7 (2013) → French classes (B2) at CIREFE univ. Rennes2 (4 hours per week)
- April 4 (2013) → English training
- May 13 to 17 (2013) → Training: Scientific Communication in English and French
- September 30 to December 20 (2013) → French course (C1) at CIREFE univ. Rennes2 (4 hours per week)

**Subjects and Seminar:**

- January 22 (2013) → Mastery of scientific information (module 1) - École VAS - Beaulieu
- 24 and 25 January (2013) → Mastery of scientific information (module 2) - École VAS - Beaulieu
- January 29 (2013) → CM: scientific publication (module 3) - École VAS - Beaulieu
- September 19 (2013) → NGS Seminar, Illumina (Rennes)

***Total of production*****Bibliographic Production**

Articles published in scientific journal – 22

Articles accepted in journals - 1

Chapters published - 3

Articles published in annals of events – 2

Presentations in Events (Congress) - 17

**Academic Advisories**

Event organization - 2

Participation in events (congress, symposium, other) - 60