Department of Food and Environmental Sciences
University of Helsinki
Finland

A MULTIFACETED STUDY OF *PROPIONIBACTERIUM FREUDENREICHII*, THE FOOD-GRADE PRODUCER OF ACTIVE VITAMIN B_{12}

Paulina Deptula

ACADEMIC DISSERTATION

To be presented, with the permission of the Faculty of Agriculture and Forestry of the University of Helsinki, for public examination in Walter Saali, Agnes Sjöberginkatu 2, on the 21st of June, at noon.

Helsinki 2017

Supervisors

Docent Pekka Varmanen
Department of Food and
Environmental Sciences
University of Helsinki
Helsinki, Finland

Docent Tuula Nyman
Proteomics Core Facility
Institute of Cinical Medicine
University of Oslo
Oslo, Norway

Docent Kirsi Savijoki
Department of Food and
Environmental Sciences
University of Helsinki
Helsinki, Finland

Professor Vieno Piironen
Department of Food and
Environmental Sciences
University of Helsinki
Helsinki, Finland

Reviewers

Doctor Soile Tynkkynen
Principal Advisor, Fresh Dairy starters
Valio Ltd, R&D
Helsinki, Finland

Doctor Veera Kainulainen
Department of Pharmacology
University of Helsinki
Helsinki, Finland

Opponent

Docent Reetta Satokari
Immunobiology Research Program
University of Helsinki
Helsinki, Finland

Custos

Professor Vieno Piironen
Department of Food and Environmental Sciences
University of Helsinki
Helsinki, Finland

Dissertationes Schola Doctoralis Scientiae Circumiectalis, Alimentariae, Biologicae

ISBN 978-951-51-3505-6 (paperback)
ISBN 978-951-51-3506-3 (PDF; http://ethesis.helsinki.fi)
ISSN 2342-5423
Helsinki University Printing House
Helsinki 2017

I humbly dedicate this work to Professor Mirosław Jarosz, who introduced me to the fascinating world of food science and beneficial microbes when I was a child and then nurtured this fascination throughout the years.

ABSTRACT

Vitamin B12 is the most complex vitamin in existence and one of the most complex non-polymeric molecules occurring in nature. It is predominantly present in animal-derived products, which places vegetarians and people with limited access to animal-derived foods at risk for developing vitamin B12 deficiency. With the current trend of limiting the consumption of foods of animal origin, the deficiency may also affect other populations.

In situ fortification of foods through microbial fermentation with food-grade bacteria is a viable method for the introduction of vitamin B12 into foods, if the microorganism is capable of synthesising the active vitamin form. Here, the capability of *Propionibacterium freudenreichii* to produce active vitamin B12 was explored with the use of a combination of microbiological and molecular approaches.

First, the activity of the heterogolously expressed and purified enzyme BluB/CobT2 was investigated. The results showed that the novel fusion enzyme was responsible for biosynthesis of 5,6-dimethylbenzimidazole (DMBI) base and its activation for attachment as the lower ligand of vitamin B12. The enzyme's inability to activate adenine, the lower ligand of pseudovitamin B12, revealed a mechanism favouring production of active vitamin B_{12} in *P. freudenreichii*. The *in vivo* study showed that formation of DMBI is oxygen dependent as no vitamin B_{12} was produced under strictly anaerobic atmosphere. Exogenous DMBI was incorporated into the vitamin molecule under both microaerobic and anaerobic conditions, with a clear preference over incorporation of adenine.

In the following study, the capability of 27 *P. freudenreichii* and 3 *Acidibacterium acidipropionici* strains to produce active vitamin B_{12} was examined by UHPLC. The yields obtained from growth in whey-based medium enriched in cobalt and supplemented with either DMBI, with the precursors of DMBI- riboflavin and nicotinamide, or without supplementation. *A. acidipropionici* strains required supplementation of DMBI to produce small amounts of active vitamin B_{12} (<0.2 μg/mL), while all of the *P. freudenreichii* strains were able to produce active vitamin B_{12} in all conditions tested. The yields of active vitamin B_{12} produced by *P. freudenreichii* and responses to supplementation were strain dependent and ranged from 0.2 to 5.3 μg/mL.

Subsequently, the active vitamin B_{12} production by the strain *P. freudenreichii* 2067 without addition of cobalt or DMBI was tested. The experiments were performed in a medium mimicking cheese environment as well as in the whey-based medium. The production of other key metabolites was examined by HPLC, while the global protein production was compared by gel-based proteomics. The results showed that regardless of different effects of the media on the metabolic state of the cells, which was reflected by distinct metabolite and protein production patterns, *P. freudenreichii* produced nutritionally relevant levels of active vitamin B12.

Finally, whole genome sequencing was employed to better characterise the species through a comparative genomics study. The use of PacBio sequencing platform, a PCR-free method producing long reads, resulted in discovery of additional circular elements: two novel, putative conjugative plasmids and three active, lysogenic bacteriophages. The long reads also permitted characterisation and classification of two distinct types of CRISPR-Cas systems. In addition, the use of PacBio sequencing platform allowed for identification of DNA modifications, which led to characterisation of Restriction-Modification systems together with their recognition motifs, many of which were reported for the first time. Genome mining suggested surface piliation in the strain *P. freudenreichii* JS18, which was confirmed by transmission electron microscopy and assessment of specific mucus binding.

Taken together, the results reported in this work support the role of *P. freudenreichii* as the only known candidate for *in situ* fortification of foods in vitamin B_{12}. The species meets the requirements of having the food-grade status and showing the capability to produce appreciable amounts of active vitamin B_{12} without the need for the addition of supplements that are forbidden in food production. The whole genome sequencing, besides contributing to a better understanding of this bacterium, will allow for the development of novel applications and facilitate further studies.

ACKNOWLEDGEMENTS

The research on which this thesis is based was carried in years 2012-2017 at the Department of Food and Environmental Sciences, University of Helsinki. The work was funded by the Academy of Finland.

I would like to use this opportunity to thank my two main supervisors, the inseparable duo Docent Pekka Varmanen and Docent Kirsi Savijoki. Individually, they each guided me in their respective areas of expertise and in their own styles: Pekka in the area of genetics, in a focused and slightly pessimistic fashion; Kirsi in the area of biochemistry and proteomics in a dynamic, creative way. Together they gave me more than can ever be asked of thesis supervisors: a sense of belonging and support, both professionally and outside the workplace. For this, I am forever grateful to you.

Secondly, I would like to thank another of my supervisors: Docent Tuula Nyman. I admire your expertise in the area of proteomics and appreciate your no-nonsense approach to supervision, but I am particularly grateful for making me feel like I belong in your tight-knit research group, even though I qualified more as a passerby.

I also thank my fourth, and most senior supervisor, also acting as Custos during my thesis defence: Professor Vieno Piironen. Thank you first of all for trusting Pekka's decision to hire me as a PhD student on your joint project but also for quickly showing trust in my abilities as well.

My thesis pre-examiners: Dr Soile Tynkkynen and Dr Veera Kainulainen. Thank you both for insightful comments on my thesis, delivered in a kind and supportive fashion.

I thank my collaborators from the Food Chemistry side, with whom we shared numerous fruitful discussions: Dr Bhawani Chamlagain, Dr Minnamari Edelmann and Docent Susanna Kariluoto. I especially thank Dr Bhawani Chamlagain who, until recently a PhD student himself, kindly shared his experience in tackling bureaucratic affairs leading to the defence.

My collaborators at the DNA Sequencing and Genomics Laboratory: Pia Laine, Dr Olli-Pekka Smolander, Dr Patrik Koskinen, Lars Paulin, and Docent Petri Auvinen. Thank you for introducing me to the vast world of sequencing and for the highly enjoyable project meetings. I take this opportunity to specifically thank Pia for helping me to make my twenty precious tables look the way I imagined them.

My most recent and most unexpected collaborator, Dr Richard J. Roberts. I am deeply honoured to be able to work with you. I never thought that someone who is at this high level in the scientific community could be so approachable and accommodating as well.

Finally, my dear friend-turned-collaborator, Dr Petri Kylli. Thank you for sacrificing your evenings and weekends (without complaint!) to run the numerous analyses showing different activities of my pet enzyme.

I owe thanks to the Microbiology and Biotechnology Doctoral Program for providing and managing my doctoral education. Specifically, thank you to Dr Kari Steffen, Dr Karen Sims-Huopaniemi and Professor Marko Virta.

Big thank you to the people who were not directly involved in my thesis work, yet affected me greatly during that time. Associate Professor Mirko Rossi, thank you for introducing me into the world of bioinformatics and for patiently answering all of the questions that followed. I would have not been able to complete the study V without your advice. Associate Professor Finn Vogensen, thank you for being my mentor and an inspiration already during my time at the University of Copenhagen, and throughout the years. Professor Emeritus Eero Puolanne, you are my role model, I hope one day I could be more like you.

This brings me to EMFOL, or the Erasmus Mundus Food of Life. Not my doctoral, but the master's program. I would like to thank the Consortium for giving me the opportunity to complete the international, double master's degree which led me to Finland and ultimately to the creation of this book.

I also thank my lovely master's students: Panchanit Sangsuwan, Maria Asuncion Fernandez Lopez, Tatiana Ischenko, Ermolaos Ververis and Subash Basnet. I think I have learnt more from you than I was able to teach you.

I would also like to thank my fellow PhD students and the researchers in Viikki: Göker Gurbuz, Dr Jose Martin Ramos Diaz, Dr Jiao Liu, Dr Xin Huang, Dr Kevin C. Deegan, Dr Anne Duplouy, Dr Rosanna Coda and the Polish crew: Dr Dorota Nawrot, Dr Wojciech Cypryk, Dr Anna Stygar, Dr Dominik Kempa, Dr Katarzyna Leskinen and Marcelina Bilicka. Dear friends, you made my time in Finland a lot brighter and more cheerful. Thank you all.

My outside-the-university friends: Aleksandra Jarosz, , Julianna Kasprzak, Anh Pham and Mikko Muhonen. Thank you for keeping me in touch with the outside world and helping my skin tone remain just a shade darker than the laboratory white.

I owe my most profound thanks to my parents, Bogusława and Bernard Deptuła. Without their love and support I would have not dared to move abroad or to pursuit a PhD. Mamo, Tato, dziękuję.

Last but not least, I would like to both thank and to apologise to my very supportive fiancé, Pasi Perkiö. I was not the easiest to be around during the thesis writing process, yet you handled it like a boss. I feel so lucky, that even after four years with you I still tend to pinch myself when I wake up. I am looking forward to what future brings us.

TABLE OF CONTENTS

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications:

I Deptula, P., Kylli, P., Chamlagain, B., Holm, L., Kostiainen, R., Piironen, V., Savijoki, K. and Varmanen, P. (2015). BluB/CobT2 fusion enzyme activity reveals mechanisms responsible for production of active form of vitamin B_{12} by *Propionibacterium freudenreichii*. Microbial cell factories, 14:.186.

II Chamlagain, B., Deptula, P., Edelmann, M., Kariluoto, S., Grattepanche, F., Lacroix, C., Varmanen, P. and Piironen, V. (2016). Effect of the lower ligand precursors on vitamin B_{12} production by food-grade Propionibacteria. LWT-Food Science and Technology, 72, pp.117-124.

III Deptula, P., Chamlagain, B., Edelmann, M., Sangsuwan, P., Nyman, T., Savijoki, K., Piironen, V. and Varmanen, P. (2017) Food-like growth conditions support production of active vitamin B_{12} by *Propionibacterium freudenreichii* 2067 without DMBI, the lower ligand base, or cobalt supplementation. Frontiers in Microbiology, 8:368

IV Koskinen, P., Deptula, P., Smolander, O.P., Tamene, F., Kammonen, J., Savijoki, K., Paulin, L., Piironen, V., Auvinen, P. and Varmanen, P. (2015). Complete genome sequence of *Propionibacterium freudenreichii* DSM 20271^T. Standards in genomic sciences, 10:83

V Deptula, P., Laine P.K., Roberts R.J., Smolander, O.P., Vihinen, H., Piironen, V., Paulin, L., Jokitalo, E., Savijoki, K., Auvinen, P. and Varmanen, P. In preparation.
De novo assembly of genomes from long sequence reads reveals uncharted territories of *Propionibacterium freudenreichii*.

The publications are referred to in the text by their roman numerals.

ABBREVIATIONS

2DE	two-dimensional gel electrophoresis
Ade-RP	7-α-D-ribofuranosyl adenosine 5′phosphate
AI	adequate intake
ALA	δ-aminolevulonic acid
ALAS	ALA synthase
AMP	9-β-d-ribofuranosyl adenosine 5′phosphate (adenosine monophosphate)
ANI	average nucleotide identities
ATP	adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
BLASTn	BLAST nucleotide
BLASTp	BLAST protein
BSA	bovine serum albumin
Cas	CRISPR-associated genes/proteins
CDD	Conserved Domains Database
CRISPR	clustered regularly interspaced palindromic repeats
DMBI	5,6-dimethylbenzimidazole
EFSA	European Food Safety Authority
ENA	European Nucleotide Archive
ESI	electrospray ionization
FMN	flavin mononucleotide
$FMNH_2$	reduced FMN
GDP	guanosine diphosphate
GMP	guanidine monophosphate
GRAS	Generally Recognised as Safe
GTP	guanosine triphosphate
HGAP	hierarchical genome-assembly process
HPLC	high performance liquid chromatograph
ICE	Integrative and Conjugal Elements
IEF	isoelectric focusing
IF	Intrinsic Factor
LB	Luria-Bertani medium
LC	liquid chromatography
MBA	microbiological bioassay
MS	mass spectrometry
NADH	nicotinamide adenine dinucleotide
NaMN	nicotinic acid mononucleotide
NCBI	National Center for Biotechnology Information
OD	optical density
ORF	open reading frame
PacBio	Pacific Biosciences
PBS	phosphate-buffered saline
PCR	Polymerase Chain Reaction
PDA	photo diode array

PIPES	1,4-Piperazinediethanesulfonic acid
PPA	propionic medium
QIT	quadrupole ion trap
QPS	Qualified Presumption of Safety
QTOF	quadrupole time-of-flight
RDA	recommended dietary allowance
rDNA	ribosomal DNA
RM	restriction-modification
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SMRT	single-molecule, real-time
SNP	single-nucleotide polymorphism
TCA	tricarboxylic acid
UHPLC	ultra-high-performance liquid chromatography
USDA	US Department of Agriculture
WBM	whey-based medium
YEL	yeast extract-lactate based medium
α-RP	α-ribazole phosphate

1 INTRODUCTION

Propionibacterium freudenreichii is an Actinobacterium that belongs to the order *Propionibacteriales*, family *Propionibacteriaceae* and genus *Propionibacterium*. The name "Propionibacterium" was first suggested in 1909 by Orla-Jensen, based on the large amounts of propionic acid produced during fermentation by these bacteria (Vorobjeva, 1999a). Over the years, the genus was reclassified multiple times, and recently, owing to the amassing genomic data, it was divided into four genera: *Propionibacterium*, *Acidipropionibacterium*, *Cutibacterium* and *Pseudopropionibacterium*. Genus *Propionibacterium* currently includes four species: *Propionibacterium freudenreichii* (type species), *Propionibacterium cyclohexanicum, Propionibacterium acidifaciens* and *Propionibacterium australiense*. The characteristics of the genus include a high GC% content, varying from 64 to 70%, and the presence of meso-2,6-diaminopimelic acid as the diagnostic amino acid in the peptidoglycan (Scholz and Killian, 2016).

Historically, and per the latest edition of Bergey's Manual of Systematic Bacteriology (Patrick and McDowell, 2012), the species *P. freudenreichii* is divided into subspecies *freudenreichii* and *shermanii*. This subdivision is based on two phenotypic characteristics: the ability to reduce nitrate and the ability to utilise lactose, with the former positive and the latter negative for the subspecies *freudenreichii* and *vice versa* for the subspecies *shermanii*. However, after genetic data became available, the subdivision was repeatedly challenged as unfounded (Dalmasso *et al.,* 2011; de *Freitas et al.,* 2015; Falentin *et al.,* 2010; Loux *et al.,* 2015; Thierry *et al.,* 2011), and the subspecies names were proposed to be considered synonyms of the species *P. freudenreichii* (Scholz and Kilian, 2016).

P. freudenreichii was originally isolated from Swiss-type cheese (van Niel, 1928), where it is currently used as a ripening culture. Its primary role in the cheese is to remove lactic acid that is produced by the lactic acid bacteria starter strains. Metabolism of lactic acid by Propionibacteria results in the production of the short-chain fatty acids propionate and acetate, as well as in the production of CO_2. Production of CO_2 is crucial for the formation of the typical, large, round eyes in Emmental cheese (Daly *et al.,* 2010). The main contribution of *P. freudenreichii* to flavour development during cheese ripening is through hydrolysis of lipids and the release of amino acids (most notably proline) and volatile compounds (Vorobjeva, 1999b). The lipolytic activity of *P. freudenreichii* accounts for over 90% of free fatty acids released during ripening and consequently plays a major role in the development of the typical flavour associated with Swiss-type cheeses (Dherbecourt *et al.,* 2010). Evidence of its long, safe use in cheeses resulted in granting the status of Generally Recognised as Safe (GRAS) to *P. freudenreichii* (Falentin *et al.,* 2010; Meile *et al.,* 2008).

Aside from its role in the dairy industry, *P. freudenreichii* is also used for the industrial production of vitamin B_{12}. Vitamin B_{12} is one of the most

complex non-polymeric molecules found in nature and the most complex vitamin. Due to the complexity and cost of the chemical synthesis of vitamin B_{12}, which requires approximately 70 steps, the vitamin is produced exclusively through microbial fermentation by some of the few synthesising species of bacteria and archaea (Martens *et al.,* 2002).

Vitamin B_{12} deficiency is associated with multiple neurological symptoms, ranging from forgetfulness and fatigue to severe and irreversible neurological disorders (Reynolds, 2006). Vitamin B_{12} is primarily found in animal-derived foods. Therefore, vegetarians and populations with limited intake of such foods are at risk for developing deficiency (EFSA NDA Panel, 2015; Pawlak *et al.,* 2013). The current trend towards introducing more sustainable diets with reduced consumption of animal-derived foods (Perignon *et al.,* 2017) is likely to put a wider population at risk for vitamin B_{12} deficiency, unless such deficiencies are corrected by supplementation (Eshel *et al.,* 2016).

In recent years, natural fortification of foods in vitamins produced by food-grade bacteria was suggested as a potential way to increase the nutritional value of food products without increasing production costs (Burgess *et al.,* 2009). At the same time, it would allow consumers to enhance their vitamin intakes from a normal diet (LeBlanc *et al.,* 2011) and eliminate the need for food supplementation with chemically synthesised vitamins (Capozzi *et al.,* 2012). *P. freudenreichii* is the only GRAS bacterium known to synthesise active vitamin B_{12} and is therefore a viable candidate to be considered for the *in situ* fortification of foods.

Despite the recognised role of *P. freudenreichii* in the dairy industry and its ability to produce vitamin B12, the bacterium remained poorly characterised on a genomic level, with only one whole genome sequence (Falentin *et al.,* 2010) publicly available for the species at the beginning of this project. The most likely reason for this is the character of the *P. freudenreichii* DNA, which has a high GC% content and stretches of repeated sequences, two known factors that hamper the sequencing (Huptas *et al.,* 2013) and assembly of bacterial genomes (Koren *et al.,* 2013), respectively.

In this thesis, the applicability of *P. freudenreichii* for *in situ* fortification of foods in active vitamin B_{12} is explored using multiple approaches. First, the activity of the BluB/CobT2 enzyme is assessed to determine its role in the final steps of the biosynthetic pathway of active vitamin B_{12} in *P. freudenreichii*. Subsequently, the roles of precursor supplementation on the increased production of B_{12} are explored. Finally, the ability of *P. freudenreichii* to produce active vitamin B_{12} without supplementation is tested in food-like growth media. Additionally, the suitability of the Pacific Biosciences (PacBio) RS II sequencing platform for overcoming the challenges posed by the DNA of *P. freudenreichii* is shown. The platform is then employed for whole genome sequencing of 18 additional strains, and a comparative genomics study is undertaken to better characterise the species on a genomic level.

2 REVIEW OF THE LITERATURE

2.1 VITAMIN B12

Vitamin B_{12} (Figure 1) belongs to a group of compounds known as cobamides, which are structurally related to compounds such as haeme and chlorophyll. The characteristic features of a cobamide include the structure of the corrin ring, which is decorated with multiple methyl, acetamide and propionamide side chains, and a cobalt ion that is coordinated in the centre of the ring. Additionally, the structure possesses two axial ligands that are coordinated to the central cobalt ion: the upper and the lower ligands. In cobalamins, which are the cobamides with vitamin B_{12} activity, the lower ligand consists of 5,6-dimethylbenzimidazole (DMBI), which is extended through an α-glycosidic link to ribose-3-phosphate (the pseudonucleotide) (Ball, 2006, Roth *et al.*, 1996). The upper ligands of cobalamins vary and include methyl, hydroxyl, adenosyl and cyano groups, resulting in the corresponding names of methylcobalamin, hydroxocobalamin, adenosylcobalamin and cyanocobalamin (Ball, 2006; Martens *et al.*, 2002).

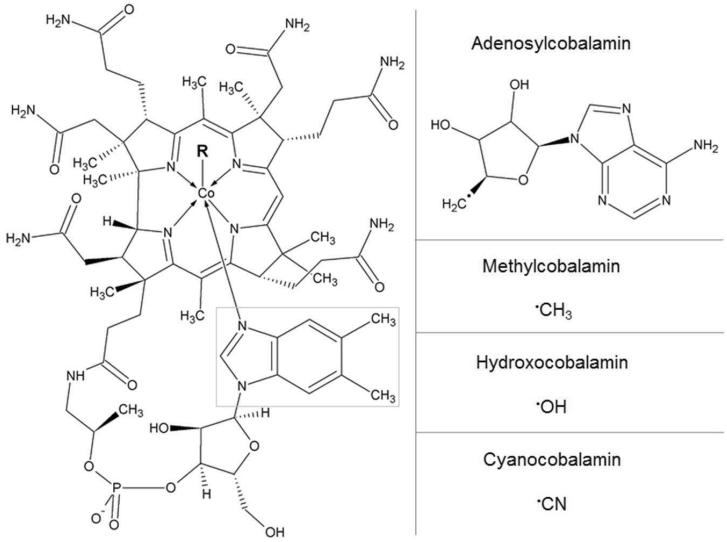

Figure 1 Structure of vitamin B_{12} with the alternative upper ligands (R) and corresponding names of complete vitamin B_{12} molecules carrying these ligands are shown on the right. The lower ligand DMBI is indicated in a box

The name vitamin B_{12} refers to the form with the cyano ligand, which does not occur in nature but is formed from other cobalamins during extraction with sodium cyanide (Battersby and Leeper, 1999). The other cobalamins are converted to the cyanocobalmin to increase stability, which protects the vitamin from degradation during heat extraction (Kumar *et al.,* 2010). Cyanocobalamin is readily converted into coenzyme forms in the human body (Obeid *et al.,* 2015). Because the upper ligand has no effect on the B_{12} activity, all of the cobalamins in this work are referred to as vitamin B_{12} for simplicity.

2.1.1 ROLE OF VITAMIN B_{12} IN HUMANS

Vitamin B_{12} was discovered by Whipple and concurrently by Minot and Murphy as a treatment for "pernicious anaemia" in 1925. Before the isolation of vitamin B_{12} in crystalline form in 1947 (Rickes *et al.*, 1948a), the treatment involved eating large portions of raw liver. Dorothy Hodgkin resolved the crystal structure of the vitamin in 1956, leading to the extensive research that enabled its further characterisation (Banerjee and Ragsdale, 2003).

In humans, the activity of vitamin B_{12} is restricted to two enzymes: cytoplasmic methionine synthase and mitochondrial methylmalonyl-CoA mutase. Methionine synthase uses methylcobalamin form of vitamin B_{12} as a co-factor in methylation of homocysteine to methionine, with concurrent demethylation of 5-methyl-tetrahydrofolate. Thus, methionine synthase plays a role in methyl-transfer reactions and in nucleotide synthesis (EFSA NDA Panel, 2015; Nielsen *et al.,* 2012; Reynolds, 2006). Methylmalonyl-CoA mutase catalyses a methylcobalamin-dependent radical-rearrangement of L-methylmalonyl-CoA to succinyl-CoA. The reaction plays an important role in the metabolism of branched-chain amino acids and odd-chain fatty acids by directing them to the tricarboxylic acid (TCA) cycle (Nielsen *et al.*, 2012; Watanabe *et al.*, 2013).

2.1.1.1 Vitamin B_{12} deficiency

In its severe form, vitamin B_{12} deficiency is associated with megaloblastic anaemia and irreversible, severe neurological disorders, whereas the milder form can present with non-specific symptoms such as lethargy, forgetfulness, infertility or depression (Reynolds, 2006). Due to the lack of specificity in the symptoms, the deficiency is diagnosed through biological markers. The markers that are measured include serum levels of methylmalonic acid and/or homocysteine (Stabler, 2013) and holo-transcobalamin II (a protein that transports vitamin B_{12} that is absorbed in the ileum to cells in the body) (Nielsen *et al.,* 2012; Pawlak *et al.,* 2013); preferably, a combination of at least two of these markers is used for diagnostics (EFSA NDA Panel, 2015).

Vitamin B_{12} deficiency can stem from either extended inadequate intake or from impaired transport and metabolism. Because vitamin B_{12} predominantly occurs in animal-derived foods, the deficiency due to the extended inadequate intake is the most common among vegetarians (EFSA NDA Panel, 2015, Pawlak *et al.*, 2013, Watanabe *et al.*, 2013), and consequently, in the infants of vegetarian mothers (Black, 2008; Green and Miller, 2013). The impaired transport and metabolism can be a result of genetic disorders, which are relatively rare (Carmel, 2008; Green and Miller, 2013), or due to Intrinsic Factor (IF)-related malabsorption (EFSA NDA Panel, 2015). IF is a glycoprotein that is produced by the parietal cells of the stomach. IF is responsible for binding to vitamin B_{12} and transporting it to the ileum, where the IF-bound vitamin B_{12} is transported into the epithelial cells (Nielsen *et al.*, 2012). IF-related malabsorption can result from its insufficient synthesis

(pernicious anaemia developing due to autoimmune disorders) or from failure of the ileal uptake (Carmel, 2008).

For the prevention of deficiency due to inadequate intake in Europe, the European Food Safety Authority (EFSA) set the adequate intake (AI) levels at 1.5 μg/day for children aged 7 months to 6 years, 2.5 μg/day for ages 7-10 years, 3.5 μg/day for ages 11-14 and 4 μg/day for ages 15 and above. For pregnant and breastfeeding women, AI was increased to 4.5 and 5 μg/day, respectively (EFSA NDA Panel, 2015). In the USA, the recommended dietary allowance (RDA) was set by the US Institute of Medicine at 2.4 μg/day for persons over 14 years of age (National Institutes of Health, 2016). In the cases of impaired transport, megadoses of vitamin B_{12} as dietary supplements are recommended because it is estimated that approximately 1-2% of the vitamin can be absorbed through passive transport (EFSA NDA Panel, 2015). It was reported that long-term daily doses of up to 5 mg of vitamin B_{12} have no adverse effects (EFSA NDA Panel, 2015). However, vitamin B_{12} and related compounds have recently been shown to play a role in modulation of the human gut microbiota (Degnan *et al.,* 2014a). To date, the specific effects of this modulation on human health remain unknown.

2.1.2 VITAMIN B_{12} IN FOODS

Vitamin B_{12} is synthesised solely by microorganisms, including species found in soil, water, and sewage and also in the rumen and intestines of animals. Therefore, in animal-derived foods, vitamin B_{12} can originate either from the animal's own digestive tract or from ingestion of vitamin B_{12}-producing microorganisms or animal tissue. In plant-based foods, vitamin B_{12} can occur naturally due to microbial contamination or, in legumes, through symbiosis with B_{12}-producing microorganisms (Ball, 2006). Alternatively, the vitamin can be introduced through fermentation by B_{12}-producing microorganisms (EFSA NDA Panel, 2015). The vitamin B_{12} content of several common foodstuffs, as reported by the US Department of Agriculture (USDA), is listed in Table 1 (USDA, 2016). The foodstuffs which have been previously suggested as sources of vitamin B_{12} for vegetarians and vegans (Watanabe *et al.,* 2013) are included in the list, and it is clear that these food products are not particularly vitamin B_{12}-rich. This, combined with the fact that some vegetarians reportedly refuse to take vitamin supplements due to various beliefs (Antony, 2003; Pawlak *et al.*, 2013), calls for alternative measures for introducing vitamin B_{12} into everyday diets.

Table 1 Vitamin B_{12} content in some foodstuffs as reported by USDA. Food products suitable for lacto-ovo-vegetarians are marked in bold, while foods suitable for vegans are additionally underlined

Foodstuff	Vitamin B_{12} (µg/100g)
Clams (steamed)	99.1
Mussels (steamed)	24.0
Breakfast cereal (B_{12} fortified)	**20.0**
Mackerel (Atlantic, cooked, dry-heat)	19.0
Swiss cheese	**3.0**
Salmon (chinook, cooked, dry-heat)	2.8
Beef (lean, plate steak, cooked, grilled)	2.1
Brie cheese	**1.7**
Rockfish (cooked, dry-heat)	1.6
Turkey (cooked, roasted)	2.0
Ham (cured, roasted)	0.7
Egg (poached)	0.6
Yoghurt (fruit, low fat)	0.5
Milk (skim)	0.4
Chicken (light meat, cooked, roasted)	0.4
Seaweed (rehydrated)	0.3
Mushrooms (brown, Italian)	0.1
Tempeh	0.08

2.1.2.1 B_{12} biosynthesis for in situ fortification of foods

In situ fortification of foods in vitamins by fermentation with food-grade bacteria has been suggested as a potential way to increase the nutritional value of food products without increasing production costs (LeBlanc *et al.*, 2011) while simultaneously eliminating the need for food supplementation with chemically synthesised vitamins (Capozzi *et al.*, 2012). In recent years multiple attempts have been made towards vitamin B_{12} production in foods (Bhushan *et al.,* 2017; De Angelis *et al.,* 2014; Edelmann *et al.,* 2016; Gu *et al.,* 2015; Molina *et al.,* 2012; Van Wyk *et al.,* 2011).

The food-grade bacteria that are considered for their ability to produce vitamin B_{12} include the GRAS status-holding Lactobacilli, among them *Lactobacillus reuteri* (Taranto *et al.*, 2003), *Lactobacillus plantarum* (Bhushan *et al.*, 2017), and *Lactobacillus rossiae* (De Angelis *et al.*, 2014) and also *P. freudenreichii* and the recommended for Qualified Presumption of Safety (QPS) (Leuschner et al., 2010; EFSA BIOHAZ Panel, 2012) *Acidipropionibacterium acidipropionici* (Parizzi *et al.,* 2012).

2.1.2.2 Distinguishing active vitamin B_{12} from other cobamides

Historically, the levels of vitamin B_{12} have been assessed using the microbiological bioassay (MBA) (Degnan *et al.,* 2014b). The MBA method is based on measuring the level of growth of a microorganism that is auxotrophic for vitamin B_{12}, such as *Lactobacillus delbrueckii* (previously *leichmannii*) ATCC 7830 (Chamlagain *et al.,* 2015; Quesada-Chanto *et al.,* 1998), where the

growth level is assumed to correspond to the measurable amount of vitamin B_{12}. Although the method is officially approved and frequently used, it is not suitable for measuring active vitamin B_{12}, as it is not capable of distinguishing between different corrinoids or other compounds that stimulate the growth of the test bacteria (Ball, 2006; Degnan *et al.,* 2014b). Over time, several methods with increased suitability have been developed, and chromatography-mass spectrometry methods are currently recommended for measuring vitamin B_{12}. These methods are able to separate and distinguish the active vitamin B_{12} from related compounds (Allen and Stabler, 2008; Chamlagain *et al.*, 2015; Degnan *et al.*, 2014b). From the perspective of food fortification, it is crucial that vitamin B_{12} is produced in its active form with DMBI as the lower ligand, as other forms have many fold lower affinities to the IF glycoprotein, making them unavailable to humans (Stupperich and Nexø, 1991).

2.2 BIOSYNTHESIS OF VITAMIN B_{12}

Chemical synthesis of vitamin B_{12} has been achieved (Eschenmoser and Wintner, 1977). However, the process is so technically challenging, and therefore so expensive, that the production of vitamin B_{12} is restricted to biosynthetic fermentation by microorganisms (Murooka *et al.*, 2005).

Biosynthesis of vitamin B_{12} can be divided into three stages: 1) biosynthesis of δ-aminolevulonic acid (ALA), 2) assembly of the corrin ring and preparation for the attachment of the lower ligand, and 3) formation and activation of the lower ligand.

2.2.1 BIOSYNTHESIS OF ALA

There are two known biosynthetic pathways for ALA: C4 and C5 (Figure 2).

In the C4 pathway, also known as the Shemin pathway, ALA is formed in a one-enzyme condensation reaction of glycine and succinyl coenzyme-A by the ALA synthase (ALAS) (Menon and Shemin, 1967; Piao *et al.,* 2004). The Shemin pathway is utilised by animals, fungi and only a few bacteria, including α-proteobacteria such as *Rhodobacter sphaeroides* (Piao *et al.,* 2004) and

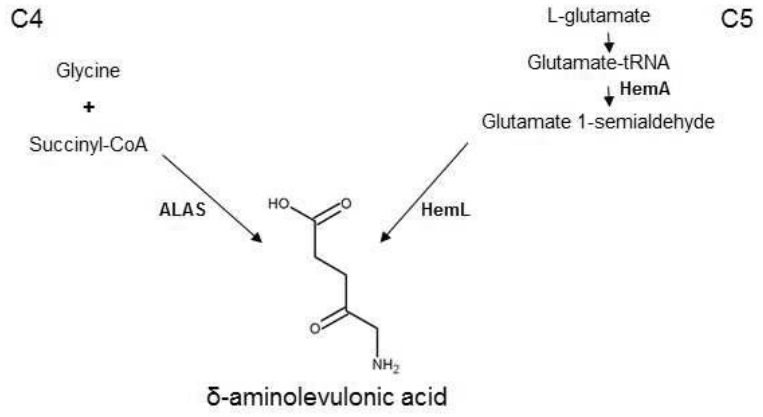

Figure 2 Biosynthesis of ALA through C4 and C5 pathways.

Rhodobacter capsulatus (Zappa *et al.*, 2010), but also in *Pseudomonas denitrificans* (Blanche *et al.*, 1995), *Streptomyces nodosus* subsp. *asukaensis* (Petříček *et al.*, 2006) and *P. freudenreichii* (Menon and Shemin, 1967).

In the C5 pathway, ALA is formed in three steps: ligation of tRNA to L-glutamate, reduction of the formed glytamyl-tRNA to glutamate 1-semialdehyde by glutamyl tRNA reductase (HemA) (Piao *et al.* 2004), and subsequent transamination by a glutamate semialdehyde aminomutase (HemL) to generate ALA (Murakami *et al.,* 1993; Piao *et al.,* 2004).

The C5 pathway is utilised by higher plants and algae (Battersby and Leeper, 1999), and it is widespread among bacteria including *Escherichia coli* (Verderber *et al.*, 1997), *Salmonella enterica* (Elliot and Roth, 1989) and *P. freudenreichii* (Hashimoto *et al.,* 1996; Oh-hama *et al.,* 1993; Piao *et al.*, 2004). Some bacteria have been reported to utilise both pathways (Avissar *et al.*, 1989; Iida and Kajiwara, 2000; Petříček *et al.*, 2006).

2.2.2 ASSEMBLY OF THE CORRIN RING

The assembly of the corrin ring proceeds from condensation of two ALA molecules into porphobilinogen; four molecules of porphobilinogen are polymerised, rearranged and cyclised to uroporphyrinogen III (Figure 3A). After formation of uroporphyrinogen III, the biosynthetic pathways of chlorophyll and haeme diverge, while the vitamin B_{12} pathway continues through methylations to precorrin-2 (Martens *et al.*, 2002; *Warren et al.,* 2002). The complete B_{12} biosynthetic pathway is present only in some species of bacteria and archaea and proceeds through two alternative routes: an oxygen-dependent and oxygen-independent route. The oxygen-dependent route has been described in detail for *P. denitrificans* (Blanche *et al.*, 1995; Warren *et al.*, 2002), and the oxygen-independent pathway was elucidated for *S. enterica* (Roth *et al.,* 1993), *P. freudenreichii* (Roessner *et al.*, 2002), and *Bacillus megaterium* (Moore *et al.,* 2013). The major differences between the pathways are the timing of the cobalt insertion into the corrin ring and the method of ring contraction (Warren *et al.*, 2002). In the oxygen-independent pathway, cobalt is inserted already into precorrin-2 (Martens *et al.*, 2002), where it assists in the ring contraction; in the oxygen-dependent pathway, the insertion of cobalt step occurs only after the ring is contracted in what is typically an oxygen-dependent manner (Warren *et al.*, 2002). With the cobalt inserted in the centre and with the corrin ring contracted, the two pathways rejoin at the formation of adenosyl-cobyric acid, to which an aminopropanol arm is added to form adenosylcobinamide (Figure 3B), which at that point is only missing the lower ligand to complete the vitamin B_{12} molecule (Martens *et al.*, 2002). For the lower ligand to be attached, the aminopropanol arm of adenosyl-cobinamide needs to be activated by guanosine triphosphate (GTP), which results in the attachment of guanosine diphosphate (GDP). Thus, activated adenosyl-cobinamide-guanosine pyrophosphate is ready for the attachment of the lower ligand, which is formed and activated in a separate pathway (Warren *et al.,* 2002).

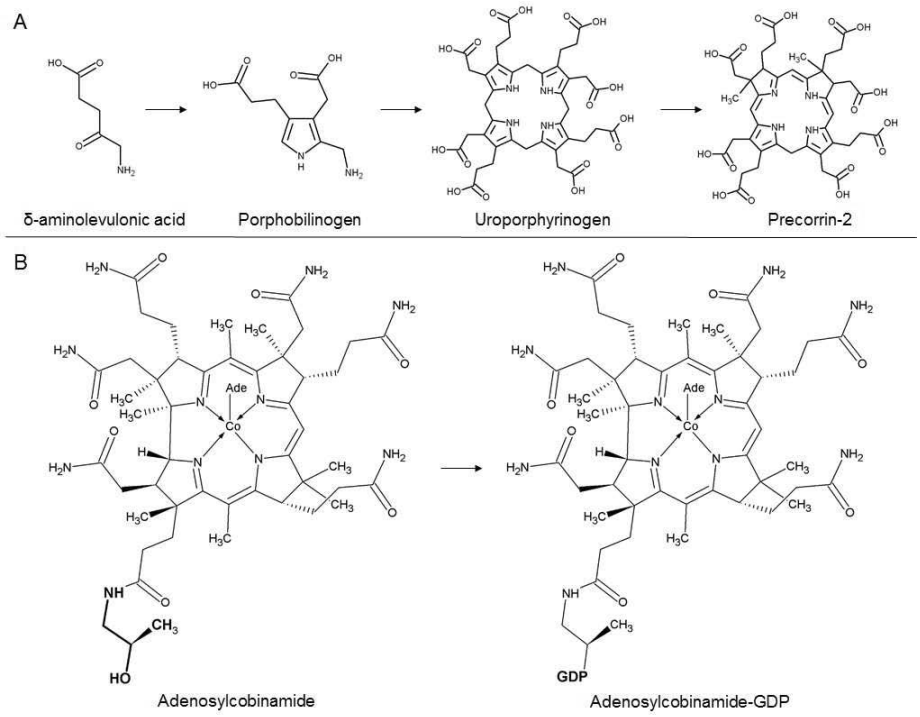

Figure 3 Formation of the corrin ring. Panel A: formation of precorrin-2 from ALA. Precorrin-2 is the first intermediate dedicated solely to the production of vitamin B_{12}. It is also the last common intermediate before the oxygen-dependent and oxygen-independent pathways diverge; Panel B: preparation of the contracted ring for attachment of the lower ligand. The aminopropanyl arm is added to adenosylcobyric acid, the first common intermediate after re-joining of the pathways, and it is marked in ***bold*** *on the molecule of the resulting adenosylcobinamide. After activation with GTP, adenosylcobinamide-GDP is ready for attachment of the lower ligand.*

2.2.3 FORMATION AND ACTIVATION OF THE LOWER LIGAND

In vitamin B_{12}, the lower ligand is formed from the DMBI base. DMBI can be synthesised in an oxygen-dependent manner from flavin mononucleotide (FMN) by bacteria that possess the *bluB* gene (Taga *et al.*, 2007; Gray *et al.*, 2007; Collins *et al.*, 2013) or in an oxygen-independent manner from glycine, glutamine, formic acid and erythrose (Renz, 1998). The oxygen-independent formation of DMBI was recently described in *Eubacterium limosum* and tied to the activity of the products of the *bzaABCDE* gene cluster (Hazra *et al.*, 2015). Prior to attachment as a lower ligand, the DMBI base needs to be activated into α-ribazole phosphate (α-RP) (Figure 4A) by the attachment of a phosphoribose moiety that is donated by nicotinic acid mononucleotide (NaMN) or a related molecule through the action of a CobT enzyme (Crofts *et al.,* 2013). Only then, adenosyl-cobinamide-guanosine pyrophosphate and

α-RP can form a vitamin B_{12} molecule (Figure 4 B), which is accompanied by the release of guanidine monophosphate (GMP).

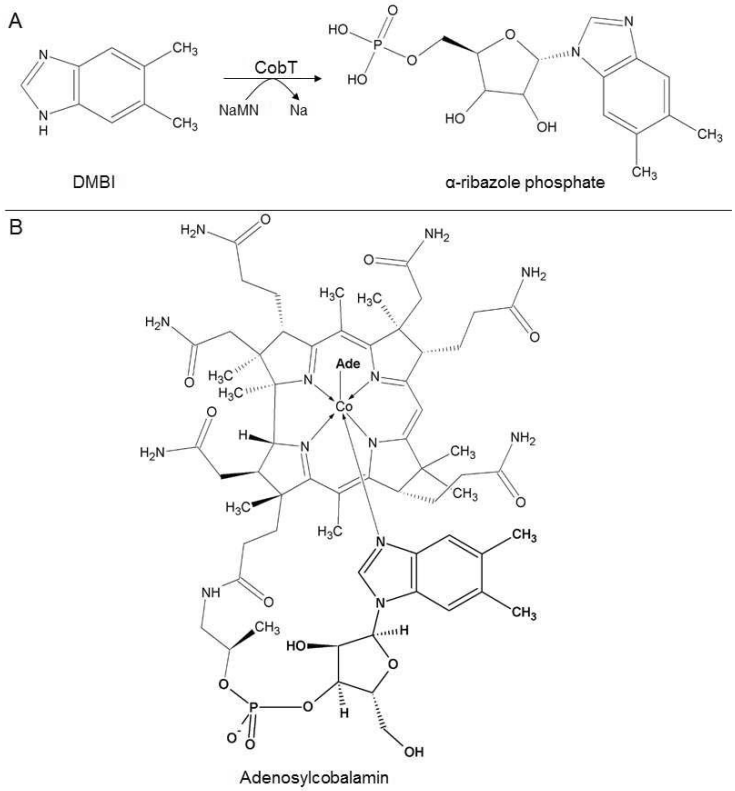

Figure 4 Final steps of the vitamin B_{12} biosynthetic pathway. Panel A: Activation of DMBI for attachment; Panel B: Complete molecule of vitamin B_{12} (in this case adenosylcobalamin). The attached lower ligand together with the pseudonucleotide appendage are marked in bold

2.2.3.1 Production of vitamin B_{12}-like compounds

It has been understood for many years that certain microorganisms can produce various cobamides, which are compounds that are structurally related to vitamin B_{12}, but can have different lower ligands. Those lower ligands include purines, various benzimidazoles, and phenolic compounds (Figure 5) (Renz, 1999; Taga *et al.,* 2008). Historically, cobamides were considered as substitutes for vitamin B_{12}, which were produced by microorganisms when DMBI was unavailable (Taga *et al.*, 2008). Specific cobamides could be synthesised in the process of so-called guided biosynthesis, in which bases for the desired lower ligand were supplied (Perlman and Barrett, 1958). While this appears to be true for some microorganisms, such as *S. enterica* (Cheong *et al.,* 2001), in some microorganisms guided biosynthesis may lead to reduced viability (Crofts *et al.*, 2013). Recent *in vitro* studies have shown that even though most the tested CobT enzymes activate DMBI preferentially (Hazra *et al.*, 2013), the activation *in vivo* depends on the preference of the

microorganism (Crofts *et al.*, 2013). The most notable example shown was that of *L. reuteri*: the CobT enzyme of *L. reuteri* activated DMBI preferentially *in vitro* (Hazra *et al.*, 2013), and also when heterologously expressed in *Sinorhizobium meliloti* (Crofts *et al.*, 2013). However, when the synthesis by *L. reuteri* was tested *in vivo,* the bacterium produced solely pseudovitamin B_{12}, a corrinoid with adenine as the lower ligand (Santos *et al.*, 2007), even when DMBI was provided (Crofts *et al.*, 2013).

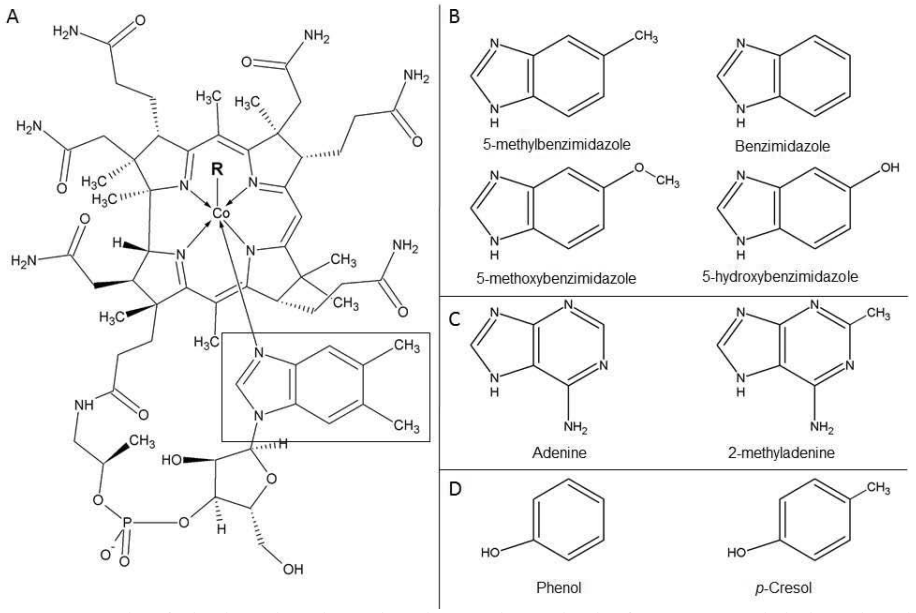

Figure 5 Examples of other lower ligands in cobamides. Panel A: Molecule of vitamin B_{12}, with the lower ligand DMBI base indicated in a box; Panel B: Benzimidazoles; Panel C: Purines; Panel D: Phenoles. The cobamide with adenine as the lower ligand is referred to as pseudovitamin B_{12}.

2.2.4 B_{12} PRODUCTION BY *P. FREUDENREICHII*

2.2.4.1 Experimental evidence

Research into the production of vitamin B_{12} by *P. freudenreichii* began shortly after the discovery of the B_{12}-production capability by microorganisms in 1948 (Rickes *et al.*, 1948b). As early as 1951, a patent (U.S. Patent 2715602) was filed by Hargrove and Leviton (1955) on a process of B_{12} production by *P. freudenreichii*. The patent stated that the yields of vitamin B_{12} depended on, aside from growth-promoting carbon and nitrogen sources, the availability of cobalt and the levels of oxygen in the culture. The microaerobic conditions were deemed the most suitable, and the maximal vitamin B_{12} yields of 0.8 µg/mL were reached. Excessive aeration was shown to result in almost complete abolition of production.

In the following patent (Speedie and Hull, 1960, U.S. Patent 2951017), a more than ten-fold increase in the vitamin B_{12} yield was obtained through the introduction of a two-step fermentation process. In this process, the cells were anaerobically grown for 70 hours, and after that time, air was allowed into the

cultures until the completion of fermentation at 168 hours. For the highest yields, culture pH was maintained at pH 7 through the addition of ammonia. The two-step incubation process was then optimised for production of B_{12} in sweet whey, which was considered a waste product of the dairy industry at the time (Bullerman and Berry, 1966a; Bullerman and Berry, 1966b; Berry and Bullerman, 1966). It was later established that to maximise yields of vitamin B_{12}, DMBI base needed to be added (Martens *et al.*, 2002; Marwaha *et al.*, 1983; Vorobjeva, 1999c), which rendered *P. freudenreichii* one of the highest natural producers of vitamin B_{12} (Martens *et al.*, 2002). In the most recent studies, the two-step incubation in the whey medium combined with supplementation with 5 mg/L of cobalt chloride and 15 mg/L of DMBI was used as the optimal method for screening of *P. freudenreichii* strains for their capability to produce vitamin B_{12} (Hugenschmidt *et al.,* 2010; Hugenschmidt *et al.,* 2011).

It should be noted that for the purposes of *in situ* food fortification, neither cobalt nor DMBI should be specifically added, as neither of these compounds has been approved for food applications (Commision Regulation No 1129, 2011). For this reason, new strategies for vitamin B_{12} production without these additions are needed.

2.2.4.2 Genetic background

The genes involved in the complete biosynthetic pathway have been deciphered for *P. freudenreichii* (Table 2). The genes are organized into 4 clusters (Murooka *et al.*, 2005; Falentin *et al.,* 2010), two of which are preceeded by B_{12} riboswitches (Figure 6), while the third riboswitch is located ahead of the *mutA* gene coding for a B_{12}-dependent methlomalonyl-CoA mutase, an enzyme involved in energy metabolism through TCA cycle (Vitreschak *et al.*, 2003). The discovery and characterisation of genes that are responsible for ALA synthesis, *hemA* and *hemL,* support the previously reported presence of the C5 pathway (Hashimoto *et al.*, 1997). Although both ALA biosynthetic pathways have been reported in *P. freudenreichii* (Iida *et al.*, 2001), no *hemA*$_s$ that encodes ALAS has been found. The studies of the genes involved in the formation of the corrin ring confirm the oxygen-independent cobalt insertion pathway (Roessner *et al.*, 2002), which is also supported by biochemical studies, regardless of the presence of oxygen (Iida *et al.,* 2007). Genome mining of the first complete genome of *P. freudenreichii* led discovery of the *bluB/cobT2* fusion gene (Falentin *et al.*, 2010). The predicted activity of the gene product points to the formation of the lower ligand DMBI from FMN in the oxygen-dependent reaction (Taga *et al.*, 2007), however this ability needs to be confirmed experimentally.

Table 2 Vitamin B_{12} biosynthetic genes identified in P. freudenreichii. The annotations and loci are derived from the whole genome sequence NC_014215.1 (Falentin et al., 2010). The organisation of the gene clusters can be seen in Figure 6

	Cluster	Gene name	Locus ID	Annotation
Cobalt transport	I	*cobA*	PFREUD_RS02240	uroporphyrinogen-III C-methyltransferase
	I	*cbiO1*	PFREUD_RS02245	cobalt ABC transporter ATP-binding protein
	I	*cbiQ*	PFREUD_RS02250	cobalt ECF transporter T component
	I	*cbiN*	PFREUD_RS11710	cobalt ABC transporter substrate-binding protein
	I	*cbiM*	PFREUD_RS02260	cobalt transporter
ALA synthesis	N/A	*hemA*	PFREUD_RS09010	glutamyl-tRNA reductase
	N/A	*hemL*	PFREUD_RS08965	glutamate-1-semialdehyde 2,1-aminomutase
Assembly and activation of the corrin ring	IV	*hemB*	PFREUD_RS08980	delta-aminolevulinic acid dehydratase
	IV	*hemC*	PFREUD_RS09005	hydroxymethylbilane synthase
	N/A	*hemD*	PFREUD_RS09540	uroporphyrinogen-III synthase
	II	*cysG*	PFREUD_RS03785	cobalamin biosynthesis protein
	II	*cbiL*	PFREUD_RS03770	precorrin-2 C(20)-methyltransferase
	II	*cobJ*	PFREUD_RS03780	tetrapyrrole methylase
	II	*cbiF*	PFREUD_RS03775	precorrin-4 C(11)-methyltransferase
	II	*cbiD*	PFREUD_RS03790	cobalt-precorrin-5B (C(1))-methyltransferase
	II	*cbiJ*	PFREUD_RS03805	cobalt-precorrin-6A reductase
	II	*cbiT*	PFREUD_RS03800	precorrin-6Y C5,15-methyltransferase (decarboxylating) subunit
	II	*cbiC*	PFREUD_RS03795	precorrin-8X methylmutase
	III	*cbiA*	PFREUD_RS05920	cobyrinic acid a,c-diamide synthase
	III	*cobA2*	PFREUD_RS05915	cob(I)alamin adenolsyltransferase/cobinamide ATP-dependent adenolsyltransferase
	III	*cbiP*	PFREUD_RS05910	cobyric acid synthase
	III	*cbiB*	PFREUD_RS05905	adenosylcobinamide-phosphate synthase
	III	*cobU*	PFREUD_RS05925	adenosylcobinamide kinase/adenosylcobinamide phosphate guanyltransferase
	III	*cobS*	PFREUD_RS05930	cobalamin synthase
	III	*cobB*	PFREUD_RS05935	TetR family transcriptional regulator
Synthesis and activation of DMBI	N/A	*bluB/ cobT2*	PFREUD_RS03140	5,6-dimethylbenzimidazole synthase

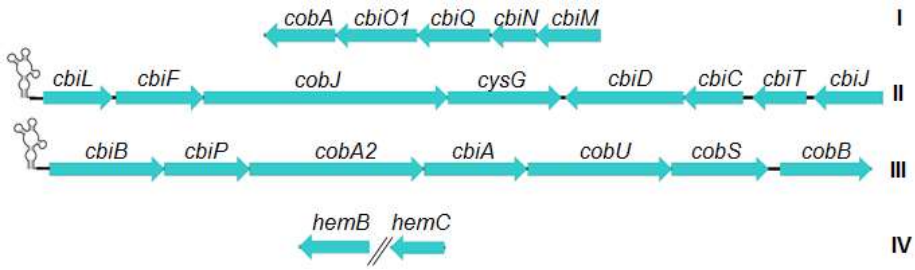

Figure 6 The four gene clusters in which majority of the vitamin B_{12} biosynthetic genes are organised in P. freudenreichii. B_{12} riboswitches can be seen upstream of genes cbiL and cbiB. Figure adapted from Falentin et al. (2010) and Polaski et al. (2016).

2.3 GENOMICS OF *P. FREUDENREICHII*

The first attempt at sequencing *P. freudenreichii* DSM 4902 (CIRM-BIA1) was publicly announced in 2004 (Meurice *et al.,* 2004). The genome sequencing was completed and announced in 2010 (Falentin *et al.,* 2010). The project shed light on multiple features of *P. freudenreichii*, including the genetic basis for the species hardiness and long-term survival. In addition, the metabolic pathways that are involved in the production of the bifidogenic factor and the formation of the propionic acid as well as other compounds that are important for flavour development in cheese were also described (Falentin *et al.,* 2010). The study also identified several misconceptions about the species. For example, the presence of all the genes that were necessary for aerobic respiration was demonstrated, even though the species was generally grown under an anaerobic or microaerobic atmosphere. Despite the species is associated with the dairy environment, it is poorly adapted to growth in milk, but adaptation to the gut environment and probiotic properties were detected. Although the strain was classified at that time as belonging to the subspecies *shermanii*, identification of genomic islands in the strain led to the first challenge of the subdivision of the species into subspecies *freudenreichii* and *shermanii* on the basis of phenotypic characteristics: utilisation of lactose and reduction of nitrate. The genes coding for β-galactosidase, the galactoside transporter and UDP-glucose isomerase were all located on an island and were most likely acquired through horizontal gene transfer, possibly from cow rumen-associated bacteria. The inability to reduce nitrate was attributed to a frameshift in the beta subunit of nitrate reductase that turned it into a pseudogene (Falentin *et al.*, 2010). The species subdivision was recently deemed not warranted (Scholz and Kilian, 2016).

In the following study, genomes of 21 strains were sequenced and assembled into draft genomes using the *P. freudenreichii* DSM4902 genome as the reference (Loux *et al.,* 2015). An additional draft genome of the probiotic strain *P. freudenreichii* ITG P20 was subsequently published (Falentin *et al.,* 2016). The draft genome data were used for various comparative and functional studies, including the genetic background of carbohydrate utilisation patterns (Loux *et al.,* 2015), adaptation to long-term survival under nutrient shortage (Aburjaile *et al.,* 2016) and a further challenge of the division into subspecies (De Freitas *et al.,* 2015).

While informative, the use of draft genomes has provided an incomplete view of the studied strains. In draft genomes information on large-scale structural rearrangements, segmental duplications and inversions or horizontal gene transfer of mobile elements is lost because of their absence from the reference genome (Chin *et al.,* 2013). The whole genome sequencing of *P. freudenreichii* was most likely restricted by the high GC% content of the DNA and the abundance of stretches of repeated sequences that hindered the sequencing and *de novo* assembly, respectively. However, the rapid development of new sequencing platforms constantly opens new sequencing possibilities.

2.3.1 SUITABLE PLATFORMS FOR THE WHOLE GENOME SEQUENCING OF *P. FREUDENREICHII*

The platform used for sequencing of the draft genomes of *P. freudenreichii*, Illumina (Loux *et al.,* 2015), relies on Polymerase Chain Reaction (PCR) amplification of the DNA sample prior to sequencing, rendering it prone to the so-called GC bias. Even though the most recent Illumina MiSeq library preparation can be conducted without the use of PCR, the performance on high GC% content genomes (≥64 GC%) remains considerably lower than for low and medium GC%-rich genomes (35-52 GC%) (Huptas *et al.,* 2016). Additionally, one of the greatest obstacles to whole genome sequencing in *P. freudenreichii* are the long stretches of sequences that are rich in repeats. The longest repeat-rich sequences found in all prokaryotic genomes are the ribosomal DNA (rDNA) operons, which range in size from 5 to 7 kbp (Huptas *et al.*, 2016) and can be found as multiple copies (Treangen *et al.*, 2009). Additional repeat sequences in prokaryotes generally arise through recombination processes such as the site-specific integration of prophages, transposition of transposable elements, gene duplication or the actions of systems such as retroelements and clustered regularly interspaced palindromic repeats (CRISPR), and the arising repeat sequences can greatly vary in size (Treangen *et al.*, 2009). It is then assumed that a read size that is capable of reading through an entire rDNA operon allows for the assembly of most microbial genomes (Koren *et al.*, 2013). To date, the longest reads produced by the Illumina MiSeq platform are cited as 2x300, with the usable quality reads placed within the range of ~150 to 190 bp (Huptas *et al.,* 2016). The Illumina's take on the short-read problem is the TruSeq protocol, in which libraries from long stretches of DNA are prepared in parallel and sequenced and assembled from short reads to create high-quality synthetic reads that reach 18.5 Kbp in size (McKoy *et al.,* 2014). The sequencing is still biased in regions where the Illumina chemistry is biased, namely, in sequences with high GC% contents and repeats (Lee *et al.,* 2016), meaning that the platform is not optimal for the sequencing of Propionibacteria.

Currently, there are two PCR-independent and long-read platforms available on the market: PacBio and MinION. PacBio was released in 2010 by Pacific Biosciences, while MinION developed by Oxford Nanopore Technologies was released in 2014 (Mikheyev and Tin, 2014).

PacBio sequencing is based on SMRT (single-molecule, real-time) technology. The elongation of the DNA sequence by an immobilised polymerase is recorded in real time as light impulses that are emitted when either of the four, differentially fluorescent-labelled nucleotides is incorporated into the molecule. The sequencing is performed on a single molecule that is made circular by the addition of hairpin adaptors and therefore sequenced multiple times (in multiple passes) in the forward and reverse orientations, without release from the active site of the polymerase. The mean read length obtained with the first- generation chemistry-C1 was in the range of approximately 1500 bp; the current chemistry-C4 has a mean of

over 20 kbp and maximum read length of 60 kbp. The additional advantage of the platform is the ability to identify DNA modification patterns, such as methylations (Rhoads and Au, 2015). The disadvantages of PacBio upon release were the high costs of the sequencer and reagents and the technology's proneness to sequencing error. Currently, several algorithms are available, such as the hierarchical genome-assembly process (HGAP) (Chin *et al.,* 2013). This algorithm, owing to the random character of the sequencing errors, with sufficient coverage improves the accuracy to more than 99.999%.

The MinION technology takes advantage of specific to each nucleotide base changes in the current, when the DNA molecule passes through a membrane-embedded protein pore. The instrument itself weighs less than 100 g, and it operates through a USB port, from which it draws power (Quick *et al.,* 2016); at its launch time, its cost was estimated at $1000 (Mikheyev and Tin, 2014). While the low accuracy and low throughput are currently an issue for MinION, the use of error-correction algorithms like the ones available for PacBio are thought to greatly improve the accuracy to approximately 99.5% (Lee *et al.,* 2016). The undeniable advantages of the instrument, aside from its low cost, are its small size and quick generation of results. These advantages enabled its use in the 2015 Ebola outbreak in remote regions of West Africa, where results were generated in under 24 hours (Quick *et al.,* 2016).

With the ongoing development of MinION, PacBio is the most suitable sequencing platform for high GC% content genomes available to date.

3 AIMS OF THE STUDY

The first aim of this study was to assess the suitability of *P. freudenreichii* for *in situ* fortification of foods in active vitamin B_{12}. To this end, it was first determined whether the predicted BluB/CobT2 enzyme could produce and activate DMBI for attachment as the lower ligand of vitamin B_{12}. Subsequently, the supplementation with riboflavin and nicotinamide, the precursors of DMBI, as substitute for supplementation with DMBI was tested. Finally, the effect of different growth media on the capability of *P. freudenreichii* to produce active vitamin B_{12} without supplementation was assessed.

The second aim of this study was to improve characterisation of the *P freudenreichii* species on the genomic level. To accomplish this, the whole genome sequencing from long sequence reads with the PacBio sequencing platform was used to sequence *P. freudenreichii* DSM 20271, one of the two type strains of the species. After determining the suitability of the method, 18 additional whole genomes of *P. freudenreichii* were sequenced and a comparative genomics analysis was performed.

4 MATERIALS AND METHODS

4.1 BACTERIAL STRAINS AND GROWTH CONDITIONS

This study was conducted on 32 bacterial strains, 27 belonging to *P. freudenreichii,* three strains of *A. acidipropionici*, one strain of *E. coli* and its one mutant (Table 3).

Table 3 Bacterial strains used in this work.

No	Strain name	Origin	Sequencing name	Other names	Species	Study
1	256	Valio Ltd.	-	-	*P. freudenreichii*	II
2	257	Valio Ltd.	JS2	-	*P. freudenreichii*	II, III, V
3	258	Valio Ltd.	-	-	*P. freudenreichii*	II, V
4	259	Valio Ltd.	JS4	-	*P. freudenreichii*	II, V
5	260	Valio Ltd.	-	-	*P. freudenreichii*	II
6	261	Valio Ltd.	JS	-	*P.freudenreichii*	II, V
7	262	Valio Ltd.	-	P2, 482	*P. freudenreichii*	II
8	263	Valio Ltd.	JS7	P4, 474	*P. freudenreichii*	II, V
9	264	Valio Ltd.	JS8	-	*P. freudenreichii*	II, V
10	265	Valio Ltd.	JS9	-	*P. freudenreichii*	II, V
11	266	cheese	JS10	-	*P. freudenreichii*	II, V
12	274	Senson Oy	JS11	E-113198	*P. freudenreichii*	II, V
13	275	Senson Oy	JS12	E-113199	*P. freudenreichii*	II, V
14	276	Senson Oy	JS13	E-113200	*P. freudenreichii*	II, V
15	277	Senson Oy	JS14	E-113201	*P. freudenreichii*	II, V
16	278	Senson Oy	-	JS278, E-113202	*A. acidipropionici*	II
17	279	Senson Oy	-	JS279, E-113203	*A. acidipropionici*	II
18	280	Senson Oy	-	JS280, E-113204	*A. acidipropionici*	II
19	281	DSM	JS15, CIRM-BIA1	DSM 4902, NCDO 853, ATCC 9614, CIP 103027	*P. freudenreichii*	I, II, V
20	282	DSM	JS16, DSM 20271	NCDO 564, ATCC 6207, CIP 103026	*P. freudenreichii*	II, IV, V
21	283	Valio Ltd.	JS17	-	*P. freudenreichii*	II, V
22	284	Valio Ltd.	JS18	P15-90	*P. freudenreichii*	II, V
23	285	Valio Ltd.	-	P190	*P. freudenreichii*	II
24	286	Valio Ltd.	JS20	-	*P. freudenreichii*	II, V
25	287	Valio Ltd.	JS21	PS145	*P. freudenreichii*	II, V
26	288	Valio Ltd.	JS22	-	*P. freudenreichii*	II, V
27	289	Valio Ltd.	JS23	-	*P. freudenreichii*	II, V
28	290	Valio Ltd.	-	-	*P. freudenreichii*	II
29	291	Valio Ltd.	JS25	-	*P. freudenreichii*	II, V
30	292	Valio Ltd.	-	-	*P. freudenreichii*	II
31	KRX	Promega		-	*E. coli*	I
32	PD3	study I		KRX pFN18A-bluB/cobT2	*E. coli*	I

Twenty *P. freudenreichii* strains were obtained from the dairy company Valio Ltd., one was isolated from a commercial cheese starter. Four strains were isolated from barley grains by the malting company Senson Oy, together with the three *A. acidipropionici* strains. The two type strains were purchased from the culture collection Deutsche Sammlung von Mikroorganismen und Zellkulturen.

The *P. freudenreichii* strains were grown in either yeast extract-lactate based medium (YEL) (transmission electron microscopy and adhesion assay, study V), its variation propionic medium (PPA) (study I, III, IV, V) or in the industrial-type whey-based medium (WBM) (study II and III.). *A. acidipropionici* strains were grown in WBM (study II). The strains were routinely cultured at 30 °C. The cultures were always started by streaking from glycerol (15%) stocks, which were maintained at -80 °C, onto solid medium and grown for 4 days in anaerobic jars (Anaerocult, Merck). Subsequently, liquid precultures were inoculated from well-separated colonies and grown for 3 days under a microaerobic atmosphere, while the experimental cultures were routinely inoculated at the level of OD_{600}= 0.05 and grown for 3-7 days under a microaerobic atmosphere.

E. coli strains were grown on Luria-Bertani medium (LB), with 100 µg/mL ampicillin when needed, at 37 °C or 25 °C, under aerobic atmosphere with shaking at 275 RPM.

4.2 PRODUCTION AND ANALYSIS OF VITAMIN B_{12}

4.2.1 GUIDED BIOSYNTHESIS (STUDY I)

To study the preference of *P. freudenreichii* DSM 4902 towards production of the active vitamin versus pseudovitamin, the cultures were supplemented with lower ligand bases DMBI (100 µM) or adenine (100 µM); control cultures received no supplementation. Three biological replicates were performed for all cultures. To ensure efficient formation of the corrin rings, 5 µg/mL cobalt chloride was added to all conditions. The cultures were incubated statically under a normal atmosphere (microaerobic condition) or in an anaerobic chamber (Don Whitley Scientific) for 7 days after which samples were harvested for analysis of cobamides.

4.2.2 SUPPLEMENTATION WITH DMBI PRECURSORS (STUDY II)

To study the capability of Propionibacteria and Acidipropionibacteria to synthesise DMBI, cultures of 27 *P. freudenreichii* and three *A. acidipropionici* (see Table 3) strains were grown in WBM medium with 5 mg/L $CoCl_2$, supplemented with DMBI precursors: riboflavin (40 mM) and nicotinamide (27 mM). For comparison, cultures with DMBI (100 mM) added after 6 days of incubation (Hugenshmidt *et al.*, 2010) and cultures with only $CoCl_2$ were used as controls. The cultures were grown for 7 days in a two-step incubation process, where for the first three days, the cultures were grown in static cultures under an anaerobic atmosphere (Anaerocult, Merck). Oxygen was

then allowed in, and the cultures were incubated for the remaining time with shaking (150 RPM; Certomat H, Sartorius) followed by harvesting for the cobamide analysis.

4.2.3 DETECTION OF VITAMIN B_{12} (STUDY I, II AND III)

For the detection of all of the vitamin B_{12}-like compounds- cobamides and cobinamide, a method described by Chamlagain *et al.* (2015) was used, with some alterations, detailed in subsequent sections.

4.2.3.1 Extraction (study I, II and III)

For the extraction of the cobamides and cobinamide, bacterial cells were harvested from liquid cultures by centrifugation and 0.1-0.2 g of the bacterial pellets were boiled with 10 mL of the extraction buffer (8.3 mM sodium hydroxide and 20.7 mM acetic acid, pH 4.5). The buffer contained 100 µL of 1% NaCN to convert all upper ligands of the cobamides to their most stable cyano forms. This was performed to prevent losses of the compounds due to heating (Kumar *et al.*, 2010).

4.2.3.2 Immunoaffinity purification (study I)

The cell extract was filtered (0.45 µm; Pall) and loaded onto the immunoaffinity column (R-Biopharma). The column was then washed with 10 mL of MilliQ water, and the retentate was eluted with 3.5 mL of methanol and evaporated under a stream of nitrogen on a 50 °C heat block. The residue was dissolved in 300 µL of MilliQ water and filtered (0.2 µm; Pall) into the LC vial (Waters).

4.2.3.3 UHPLC-PDA (study I, II, III)

For the analysis of the cobamides and cobinamide, the ultra-high-performance liquid chromatography (UHPLC) method was used as previously described (Chamlagain *et al.,* 2015). Briefly, a Waters UPLC® system (Waters) with a C18 column (Waters Acquity HSS T3, 2.1 × 100 mm, 1.8 µm) at a flow rate of 0.3 mL/min and a UV detection by a photo diode array (PDA) detector at 361 nm was used. The injection volume was 10 µL. B_{12}-like compounds were identified based on their retention times and their absorption spectra and quantified against the cyanocobalamin standard. In studies II and III, the yield of the detected cobamides was expressed in two ways: weight/volume of the growth medium and weight/weight of the wet cell mass.

4.2.3.4 Identification (study I and II)

To confirm the identity of the cobamides and cobinamide, the extracts were analysed using a high-resolution quadrupole time-of-flight (QTOF) mass spectrometry (MS) (Synapt G2-Si; Waters) with an electrospray ionisation (ESI) interface to the UHPLC system (study I) or an Esquire-LC quadrupole ion trap (QIT) mass spectrometer with an ESI interface (Bruker Daltonics) in positive ion mode (study II), as previously described (Chamlagain *et al.*, 2015).

4.2.4 ACTIVITY OF THE BLUB/COBT2 FUSION ENZYME (STUDY I)

4.2.4.1 BluB/CobT2 activity reactions (study I)

The reactions were carried out in 50 µL volumes, containing 5 µM of the enzyme, in triplicate. The concentrations of the tested compounds were 100 µM of FMN with 0, 20 or 40 mM of nicotinamide adenine dinucleotide (NADH), 100 µM DMBI, 100 µM adenine, and 200 µM NaMN. The reactions were incubated at room temperature and protected from direct light. For the reactions containing FMN, dark tubes were used. After the indicated incubation time, the reactions were stopped by addition of 6.5% trichloroacetic acid. Products were analysed by UHPLC-ESI-MS method. For each reaction, a control without the enzyme addition was included.

4.2.4.2 UHPLC-ESI-MS method for BluB/CobT2 activity reactions (study I)

UHPLC-ESI-MS is detailed in study I. Briefly, the system consisted of a Waters Acquity UPLC® I class binary solvent manager, sample manager and column thermostat (25 °C), a C18 column (Waters Acquity UPLC® BEH, 2.1 x 50 mm, 1.7 µm) at a 0.3 mL/min flow rate and a Waters Xevo TQ-S triple quadrupole mass spectrometer (Waters). The injection volume was 1 µL. DMBI and adenine were quantified against authentic standards, while α-RP and 7-α-D-ribofuranosyl adenosine 5′phosphate (Ade-RP) were quantified using DMBI and adenine as reference standards because authentic standard compounds were lacking.

4.3 2DE AND PEPTIDE IDENTIFICATION (STUDY III)

The comparative proteome analyses of the *P. freudenreichii* 2067 (see Table 3) cells were performed using two-dimensional gel electrophoresis (2DE). Cells were harvested at two growth stages—the mid-exponential growth phase and early stationary phase as determined by the growth curve.

4.3.1 GROWTH CURVE AND SAMPLING TIMES (STUDY III)

To determine the time needed for cultures grown in the PPA and WBM media to reach the mid-exponential growth phase, a growth curve was drawn by measuring the optical density (OD) at 600 nm (UV-1800 Shimadzu, Ordior) every 4 hours to estimate the time until the mid-exponential growth phase was reached (30 hours for PPA and 24 hours for WBM). Subsequently, from the experimental cultures grown in the two-step incubation process (as described in the previous section), the samples were taken at the mid-exponential growth phase, at the early stationary phase (72 hours), and after 7 days (168 hours). The cell samples that were harvested at the first two time points were subjected to protein extraction, purification and 2DE, while the cobamides from the cells and sugars and acids from the supernatants were analysed at all three time points (see below).

4.3.2 2DE (STUDY III)

2DE analyses of *P. freudenreichii* cells cultured in PPA and WBM were done using six biological replicates. The cells were washed with ice-cold 50 mM Tris-HCl (pH 8.0), disrupted by beating with 0.1 mm glass beads (Sigma Aldrich) using Fast-Prep™-24 (MP Biomedical) in three 30-s lysis cycles at 6,5 m/s and separated by cooling for 1 minute on ice. The samples were purified with the 2-D Clean-Up kit (GE Healthcare), and the proteins were quantified using the 2D Quant kit (GE Healthcare). Fifty micrograms of protein from each sample were solubilised in UTCT buffer. Solubilised proteins were separated in the first dimension by isoelectric focusing (IEF) (Ettan IPG phore, GE Healthcare) with IPG strips (11 cm, pH 4-7; GE Healthcare) and rehydrated overnight in De-Streak solution (GE Healthcare) with 1% IPG buffer pH 4-7 (GE Healthcare). After the IEF programme completed, the strips were equilibrated in equilibration buffers and loaded onto 12.5-% Tris-HCl Criterion Precast Gels (BioRad). Separation in the second dimension was run using the Criterion Dodeca Cell (BioRad) in Tris-Glycine-SDS buffer (BioRad). After electrophoresis, the gels were stained with the SYPRO Orange dye (Sigma Aldrich) (1:5000) using a modified protocol described by Malone *et al.* (2001). The images were acquired with the Alpha Imager HP documentation system (ProteinSimple). 2DE images were analysed using the SameSpots software v4.5 (Totallab). The spots showing fold-changes ≥1.4-1.5 between the two conditions (PPA vs. WBM) and with ANOVA $p < 0.05$ were picked for the in-gel tryptic digestion and LC-MS/MS analysis.

4.3.3 IDENTIFICATION OF PEPTIDES (STUDY III)

The fluorescent 2DE gels were restained with silver nitrate using the mass spectrometry-compatible method described by O'Connell and Stults (1997), with the acetic acid concentration reduced to 0.5% in the fixing step. The selected spots were subjected to trypsin digestion (Shevchenko *et al.*, 1996) followed by nano-LC-MS/MS identification (Ultimate 3000 nano-LC, Dionex) and QSTAR Elite hybrid quadrupole TOF mass spectrometry (Applied Biosystems / MDS Sciex with nano-ESI ionisation, Sciex), as previously described (Öhman *et al.*, 2010).
The nano-LC-MS/MS data were searched through the ProteinPilotTM3.0 interface (version 2.0.1, Applied Biosystems/MDS SCIEX) using the Mascot 2.2.03 search algorithm (Matrix Science) against the NCBI database with taxonomy set to "Other Actinobacteria (class)". A successful identification cut-off was set to a significant match ($p < 0.05$) with at least two peptide hits (ion score>40).

4.4 GENETIC METHODS

4.4.1 PCR (STUDY I AND V)

PCR was used for the amplification of the *bluB/cobT2* coding region from *P. freudenreichii* DSM 4049 and for detection of the *bluB/cobT2* gene in the

pFN18A-bluB/cobT2 expression vector and in the *E. coli* PD3 strain (study I). Furthermore, PCR was employed for detection of the bacteriophage in the strain JS7 (study V) and confirmation of the duplicated region in strain JS17 (study V). All PCRs were prepared with the Phusion Mastermix (ThermoFisher Scientific), and the primers were ordered from Oligomer Oy. For the amplification of DNA from *P. freudenreichii,* 0.3% of DMSO (ThermoFisher Scientific) was used as instructed for high GC% content DNA. After being resolved on 0.8% agarose (BioRad), the gels were stained with ethidium bromide (0.5 µg/mL) (Sigma-Aldrich) and the results were visualised using the Alpha Imager documentation system (ProteinSimple).

4.4.2 ISOLATION OF GENOMIC DNA (STUDY I, IV, V)

For the DNA extraction, the Propionibacteria cells from 72 hour cultures were harvested by centrifugation for 5 minutes at 21000 g at 4 °C and washed once with 0.1 M TRIS-HCl pH 8.0. The DNA extraction was performed with the ILLUSTRA™ bacteria genomicPrep Mini Spin Kit (GE Healthcare) with 10 mg of lysozyme in the lysis buffer and a duration of 30 minutes for the lysis step. DNA purity was assessed by spectroscopic measurements with NanoDrop 1000 (ThermoFisher Scientific).

4.4.3 CLONING AND EXPRESSION OF *BLUB/COBT2* (STUDY I)

The *bluB/cobT2* coding region (NC_014215.1) was PCR-amplified from *P. freudenreichii* DNA using primers designed to append recognition sites for *SgfI* and *PmeI* restriction enzymes to the protein-coding region. The amplified product was purified with the GenElute™ PCR Clean-Up Kit (Sigma Aldrich). The created insert and the Flexi® Vector pFN18A (Promega) were treated with the Flexi Enzyme blend (*SgfI* & *PmeI*) (Promega) at 37 °C for 2 h. After a second purification step, the insert was ligated to the vector with T4 DNA ligase M0202S (New England BioLabs). The Single Step KRX Competent Cells (Promega) of the *E. coli* KRX strain were transformed with 25 ng of the resulting pFN18A-bluB/cobT2 expression vector and cultured on LB agar with 100 µg/mL ampicillin. The presence of the *bluB/cobT2* coding sequence in the clones was confirmed by whole-cell PCR using *bluB/cobT2*-specific primers. The vector from selected clone *E. coli* PD3 was purified with the GenElute Plasmid MiniPrep kit (Sigma Aldrich) and sequenced at the Institute of Biotechnology, University of Helsinki, to confirm the correct structure of the vector.

BluB/CobT2 protein overexpression was induced by culturing with 0.1% rhamnose. Protein purification was performed per the HaloTag Protein Purification System manual (Promega). The expression, solubility and purity of the recombinant protein was confirmed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The enzyme concentration was determined with NanoDrop 1000 (ThermoFisher Scientific), and the protein was stored in aliquots at -20 °C until use.

4.4.4 WHOLE GENOME SEQUENCING (STUDY IV AND V)

The complete finished genome sequences of 19 *P. freudenreichii* strains were generated at the Institute of Biotechnology, University of Helsinki, with the PacBio RS II sequencing platform (PacificBiosciences) using two SMRT cells per strain and a 120-min (chemistry P4/C2) or 240-min (chemistry P5/C3) movie time. For the *de novo* assembly of the genomes, HGAP V2 was implemented in SMRT Analysis package v.2.1.0 using default parameters, with the expected genome size and seed cut-off parameters set to 2,700,000 and 7,000, respectively (study IV), or the HGAP V3 in SMRT Analysis package v.2.3.0 was implemented using default parameters, with the genome size estimate set to 3,000,000 bp (study V). Genes were identified using Prodigal v2.50 (study IV) or Prodigal v. 2.6.2 (study V) with manual curation in the ARGO Genome Browser and submitted to the NCBI (study IV) or European Nucleotide Archive (ENA) (study V).

4.5 BIOINFORMATICS ANALYSES

4.5.1 SEQUENCE ALIGNMENTS (STUDY I, III AND V)

Sequence alignments were performed with the BLAST suite on the website of the National Center for Biotechnology Information (NCBI). For the newly sequenced genome queries, a stand-alone version of the Basic Local Alignment Search Tool (BLAST) software on the Biolinux 8 workstation (Field *et al.* 2006) was used (study V). For multiple sequence alignments, ClustalW (study I) or Clustal Omega (study V) were used.

4.5.2 COMPARATIVE GENOMICS (STUDY V)

For the comparative genomics analysis of the 20 genomes, GFF3 annotation files were generated by PROKKA (Seemann 2014). These files were then used as input for Roary (Page *et al.* 2015) to estimate the core genome and pangenome and to identify genes that were unique to each strain.

4.5.3 GENOME CHARACTERISATIONS (STUDY V)

Average nucleotide identities (ANI) were calculated using JSpecies V1.2.1 (Richer and Rosselló-Móra 2009). The EMBL files obtained from ENA were converted to the GenBank format with Seqret, which is a part of the EMBOSS package on the BioLinux 8 workstation (Field *et al.,* 2006). The obtained files were used for visualisation of the genome organisations with the Progressive Mauve algorithm in the Mauve alignment tool (Darling *et al.,* 2010). The genomic islands were detected using the IslandViewer 3 software (Dhillon *et al.,* 2015). Prophages were predicted with the Prophinder (Lima-Mendez *et al.*, 2008) and Phaster (Arndt *et al.*, 2016) online software and visually reviewed for structural completeness. CRISPR loci were identified by CRISPRFinder (Grissa *et al.,* 2007) and manually reviewed for co-localisation with the CRISPR-associated (Cas) genes. The immunity to known bacteriophages was tested by searching the spacer sequences against NCBI-tailed bacteriophages

and the whole nucleotide collection excluding Propionibacteria (txid1743) using the BLAST nucleotide (BLASTn) suite (Zhang *et al.,* 2000). The restriction-modification (RM) systems were identified with REBASE (Roberts *et al.,* 2015) and curated manually by Dr Richard J. Roberts. An automatic pilus cluster search was performed using LOCP v. 1.0.0 (Plyusnin *et al.*, 2009), followed by a local BLAST protein (BLASTp) (Altschul *et al.,* 1997) search of all genomes. The results were visualised with iTOL (Letunic and Bork, 2016), Phandango 0.8.5 (Hadfield *et al.,* 2017), EasyFig (Sullivan *et al.*, 2011) and PigeonCad (Bhatia and Densmore 2013).

4.6 OTHER METHODS

4.6.1 TRANSMISSION ELECTRON MICROSCOPY (STUDY V)

Negative staining transmission microscopy was employed for the strains in which pilus clusters were identified by the LOCP (see below). The strains were grown on YEL agar for 7 days. Single colonies were picked and suspended in 0.1 M 1,4-piperazinediethanesulfonic acid (PIPES) buffer, pH 6.8. A 3 µl aliquot was added to Pioloform-coated 200-mesh copper grids (Emitech K100X, Emitech Ltd.) and incubated for 1 minute. After that time, excess of the suspension was soaked off by a filter paper and the grids were negatively stained with 1% neutral uranyl acetate for 15 seconds. After air-drying the stained copper grids, the images were acquired at 120 kV with a Jeol JEM-1400 microscope (Jeol) using an Orius SC 1000B CCD-camera (Gatan).

4.6.2 ADHESION ASSAY (STUDY V)

Adhesion of the *P. freudenreichii* strains (JS, JS16, JS18, JS20 and JS22) on immobilized porcine mucin (Sigma-Aldrich) in 96-well Polysorp microplates (Nunc Immuno plates, Nunc) was conducted according to Terraf *et al.,* (2014) with modifications. The plates were covered with 300 µl of 0.2 mg/mL mucin in phosphate-buffered saline (PBS) with pH 7.5 (Thermo Fischer Scientific) and incubated for 30 min at 37 °C with shaking at 250 rpm (PST-60HL Thermo-Shaker, Biosan) and then statically overnight at 4 °C. Wells were washed twice with PBS and blocked with 1% bovine serum albumin (BSA) in PBS, at room temperature. Replica microplates treated with PBS without mucin and then with 1% BSA as above were also prepared to exclude BSA-specific binding. After 2 hours of incubation, mucin- and BSA-coated wells were washed twice with PBS and allowed to air-dry. For adhesion, 200 µl of cells suspended in PBS (OD_{600}=2.0) from each strain was added into the mucus- and BSA-coated wells, and incubated at 37 °C for 2 hours with shaking at 250 rpm. PBS alone was included in all experiments to subtract mucin binding by PBS. Non-adherent cells were removed and wells were washed with PBS. Then, adherent cells were stained with 200 µL of the crystal violet solution (0.1%, w/v) (Sigma–Aldrich, Munich, Germany) for 30 min at room temperature. Excess staining was washed with deionized water and the stained cells were suspended in 30% acetic acid by shaking at 400 rpm at room

temperature, and recorded at 540 nm using an ELISA reader (Labsystems Multiskan EX). Up to four independent experiments were performed, each with at least sixteen technical replicates.

Differences between sample and PBS control means were determined by one-sample t-test using statistical package program (IBM SPSS Statistics v24 for Windows, IBM). P values <0.05 were considered statistically significant.

4.6.3 ANALYSIS OF LACTOSE AND ACIDS (STUDY II AND III)

Lactose and acids (lactic acid, propionic acid, acetic acid, pyruvic acid and succinic acid) were analysed from the growth medium after removal of the cells by centrifugation and filtering (0.45 mm; Pall, USA). The analysis was performed as reported by Hugenschmidt *et al.* (2010), with an high-performance liquid chromatography (HPLC) system equipped with a pump (Waters 515), an autosampler, a UV detector set at 210 nm (Waters 717) for acids and a refractive index detector (HP 1047A) for lactose run on an Aminex HPX-87H column (7.8 300 mm, 9 mm particles; BioRad) at flow rate of 0.6 mL/min with the injection volume of 40 μL. Quantification was based on calibration curves of true standards. In study III, the molar ratios between the propionic and acetic acids produced was calculated.

5 RESULTS AND DISCUSSION

5.1 PRODUCTION OF ACTIVE VITAMIN B_{12}

P. freudenreichii is a known producer of vitamin B_{12} and it is used for industrial production of the compound (Martens *et al.,* 2002). However, for production at the industrial scale, both cobalt and DMBI are added to overcome production bottlenecks. Neither of these compounds can be added to food (Commision Regulation No 1129, 2011). Therefore, in studies I, II and III the ability of *P. freudenreichii* to synthesise active vitamin B_{12} with and without the addition of cobalt, DMBI, and other supplements was explored.

5.1.1 ACTIVITY OF THE FUSION ENZYME BLUB/COBT2 (STUDY I)

The genes involved in vitamin B_{12} biosynthetic pathway have been described for *P. freudenreichii* (Murooka *et al.,* 2005; Murakami *et al.,* 1993; Roessner *et al.,* 2002; Sattler *et al.,* 1995), except for the stage where the lower ligand base, DMBI, is synthesised and activated for attachment as a lower ligand of the active vitamin B_{12}. With the publication of the first complete genome sequence of *P. freudenreichii* DSM 4049 (Falentin *et al.*, 2010), a *bluB/cobT2* fusion gene coding for a putative phosphoribosyl transferase/nitroreductase fusion enzyme BluB/CobT2 was discovered. The predicted activities of DMBI synthase and nicotinate-nucleotide-DMBI phosphoribosyl transferase suggested that *P. freudenreichii* BluB/CobT2 is capable of synthesising and activating DMBI within one enzyme structure, a phenomenon not reported before.

To assess the activity of the fusion enzyme, the *bluB/cobT2* gene from one of the type strains, *P. freudenreichii* DSM 4049 (JS15) was cloned and overexpressed it in *E. coli* KRX using the rhamnose-inducible expression and purification system, HaloTag® (Promega). Subsequently, *in vitro* studies with the purified enzyme to assess the activity of BluB as a DMBI synthase, along with sequential DMBI synthesis and activation by BluB/CobT2 were performed and the specificity of CobT2 towards DMBI was tested. Finally, guided biosynthesis of cobamides was performed *in vivo*.

5.1.1.1 The DMBI synthase activity of BluB

BluB is responsible for the formation of DMBI from reduced FMN ($FMNH_2$) in presence of oxygen (Taga *et al.*, 2007). To reduce FMN to $FMNH_2$, a chemical reduction with NADH (Collins *et al.*, 2013) at 0, 20 and 40 mM was employed. The analysis of the results by the UHPLC-ESI-MS showed that a compound with the expected m/z ratio for DMBI (m/z 147.3) was observed at the retention time of 4.05 minutes, which was similar to the true DMBI standard (Figure 7). This confirmed the activity of BluB as a DMBI synthase.

Additionally, only trace amounts of DMBI were detected in the reaction without NADH, confirming that $FMNH_2$ was the true substrate for this reaction.

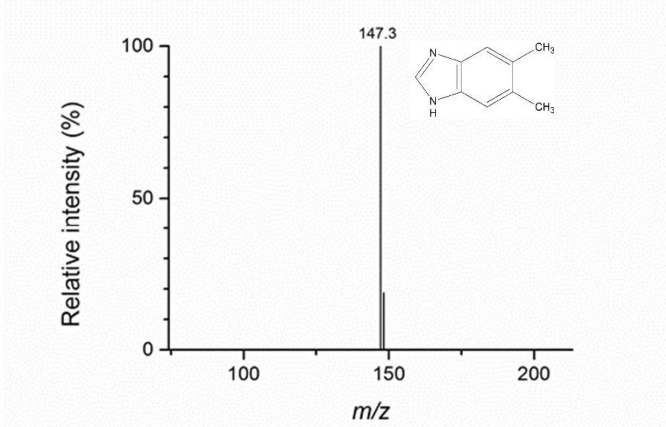

Figure 7 The MS spectrum of the product of the reaction of FMN with NADH eluting at 4.05 min. The spectrum is identical to that of the DMBI standard.

5.1.1.2 Sequential DMBI synthesis and activation by BluB/CobT2

After confirming that BluB produced DMBI from $FMNH_2$, a sequential reaction containing substrates that were necessary for the synthesis of DMBI and its subsequent activation into α-RP, namely FMN, NADH and NaMN was performed. The LC-ESI-MS analysis of the reaction products revealed a peak with a retention time of 2.75 minutes at m/z 359.1, which corresponded to α-RP. Because no true standard is commercially available for α-RP, the MS/MS fragmentation of the ion at m/z 359.1 was screened and it showed a fragment of m/z 147.3 corresponding to DMBI (Figure 8). In the sequential reaction, DMBI itself was observed only in trace amounts, even when the highest NADH concentration (40 mM) was used, suggesting an efficient activation of the produced DMBI into α-RP by BluB/CobT2 (Table 4).

Table 4 The product concentrations of the BluB/CobT2 reactions with FMN and FMN with NaMN at different concentrations of NADH. The increase of the product yield with increasing levels of NADH indicate that the reduction of FMN was a limiting factor in the reaction. The higher product yields in the sequential reaction and the lack of unactivated DMBI point to efficiency benefits of the fusion between BluB and CobT2 enzymes.

	FMN (100μM)	***FMN (100 μM) + NAMN (200 μM)***	
NADH (mM)	**DMBI (nM)**	**α-RP (nM)**	**DMBI (nM)**
0	120.19 ±6.43	-	-
20	1361.61 ±11.67	4458.41 ±200.22	0.54 ±0.18
40	2300.71 ±8.73	5094.84 ±214.89	1.72 ±0.61

Generally, the fusion of two activities in a single polypeptide may be beneficial for the enzymatic activity (Ostermeier *et al.*, 2000). In current experiments, the molar concentrations of α-RP obtained from the sequential

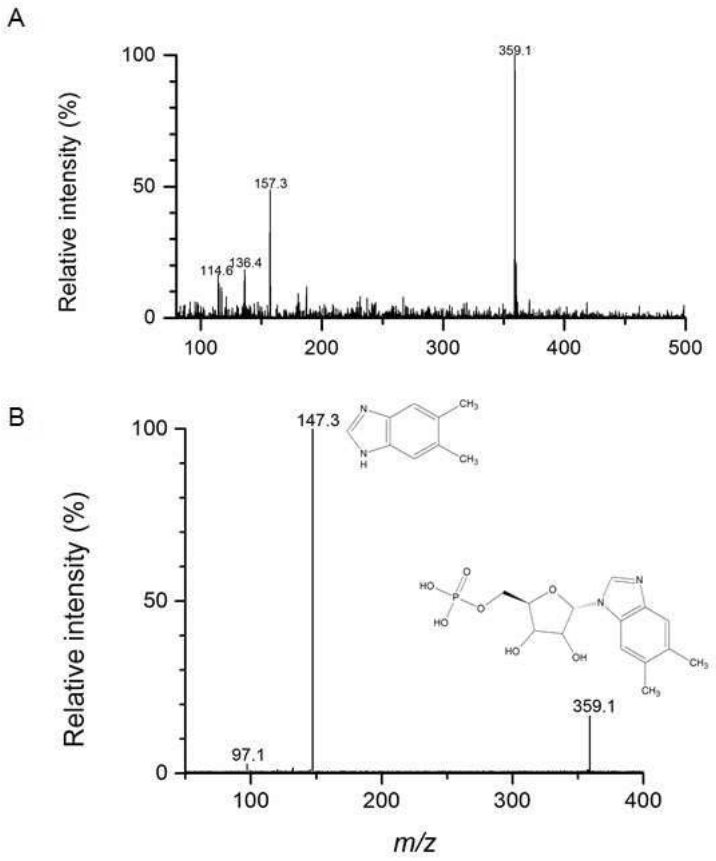

Figure 8 The product of the sequential reaction of the BluB/CobT2 enzyme with FMN, NADH and NAMn supplied. Panel A: the MS spectrum of the product eluting at 2.75 minutes and expecte m/z ratio for α-RP; Panel B: the MS/MS fragmentation spectrum of the product. The molecules of DMBI and α-RP are presented next to the ions with their corresponding mass to charge ratios (m/z).

reaction were ~2–3-fold higher than the DMBI concentrations obtained via the activity of the BluB reaction alone. It has been previously noted that benzimidazoles are only used as the lower ligands of cobamides (Crofts *et al.,* 2013). The fusion of BluB with CobT2 allows for the efficient use of the produced DMBI, which is not required for any other processes in the cell, and perhaps positively affects the substrate interaction with the enzyme.

5.1.1.3 Specificity of CobT2

From the perspective of vitamin B_{12} biosynthesis, activation of adenine is of particular interest, because adenine activated into Ade-RP would serve as a substrate for the lower ligand of the inactive, pseudovitamin B_{12}. Previous studies with semi-purified cell extracts of *P. freudenreichii* suggested that in this bacterium CobT-like activities do not include activation of adenine (Friedmann 1965; Hörig and Renz, 1980). Conversely, several of the CobT homologues studied to date have the ability to activate more than one type of

base into a corresponding α-riboside phosphate, including Ade-RP (Hazra *et al.,* 2013).

For this reason, the ability of the CobT2 enzyme to activate adenine under the same conditions that were suitable for the activation of DMBI was tested. The UHPLC-ESI-MS analysis of the reaction products resulted in a peak with a retention time of 1.25 minutes at *m/z* 348, corresponding to Ade-RP. However, the same peak with a comparable area was detected in reactions with the enzyme alone, where no adenine was added. This suggested that the observed peak corresponded to 9-β-d-ribofuranosyl adenosine 5′phosphate (AMP) rather than to Ade-RP. AMP and Ade-RP are isomers, therefore they cannot be distinguished by UHPLC-ESI-MS (Figure 9). The purification protocol used for the isolation of the enzyme involved the addition of adenosine triphosphate (ATP), which explains the presence of AMP in enzyme preparations. The fact that addition of adenine to the reaction did not increase the size of the Ade-RP/AMP peak allows conclusion that BluB/CobT2 does not react with adenine under the conditions tested.

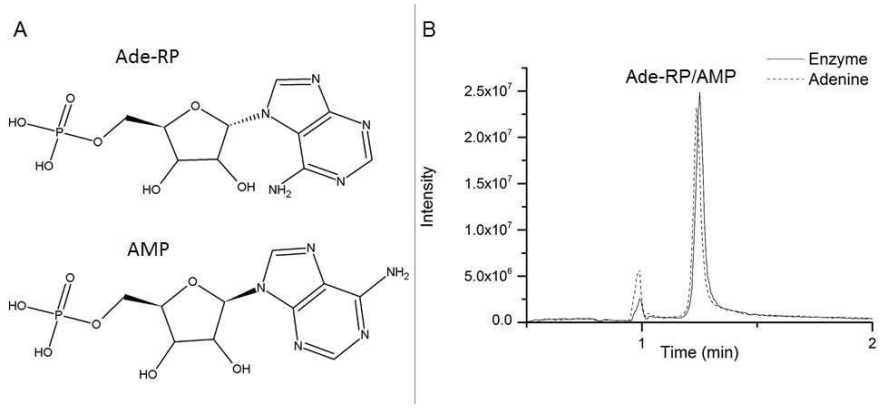

Figure 9 BluB/CobT2 does not activate adenine into Ade-RP. Panel A: the structures of the isomeric Ade-RP and AMP. The compounds differ by the attachment of adenine through 7-α-N-glycosidic bond and 9-β-N-glycosidic bond, respectively; Panel B: the LC chromatogram of the reaction with adenine and NaMN (Adenine) shows no increase in the peak corresponding to Ade-RP/AMP in relation of the enzyme preparation without reaction substrates (Enzyme).

5.1.1.4 Guided biosynthesis tests in vivo

It has been previously shown that the preference of a microorganism towards producing a specific cobamide *in vivo* can differ from the CobT activity observed *in vitro*, as in the case of the *L. reuteri* CobT enzyme. Based on the sequence analysis, the enzyme was predicted to activate both DMBI and adenine, which was confirmed by *in vitro* studies. Additionally, the enzyme activated DMBI when heterologously expressed in *S. meliloti*, a producer of active vitamin B_{12}. However, *in vivo L. reuteri* produced only the pseudovitamin (Crofts *et al.*, 2013).

To confirm the *in vitro* preference of *P. freudenreichii* CobT2 to activate DMBI, guided biosynthesis tests *in vivo* were performed. Briefly, the strain

was supplied with DMBI and adenine and cultivated under strictly anaerobic and microaerobic conditions. After 7 days, the cobamides that were produced were analysed. The results showed that under the strictly anaerobic atmosphere, no active vitamin B_{12} was made unless exogenous DMBI was added, confirming that oxygen is necessary for the formation of DMBI in *P. freudenreichii* (Figure 10).

In the cultures where exogenous DMBI was supplied, only active vitamin B_{12} and no pseudovitamin was detected. However, in cultures grown under the microaerobic atmosphere without exogenous DMBI, both cobalamin and pseudocobalamin were detected, suggesting that DMBI synthesis was a limiting factor in the production of vitamin B_{12} under the growth conditions used. Conversely, supplementing the cultures with adenine did not increase the levels of pseudovitamin, even under anaerobic conditions, but instead cobinamide, a cobamide missing the lower ligand, was detected. These results confirm that *P. freudenreichii* preferentially produces the active form of vitamin B_{12} and that the production of DMBI is the limiting factor in this process.

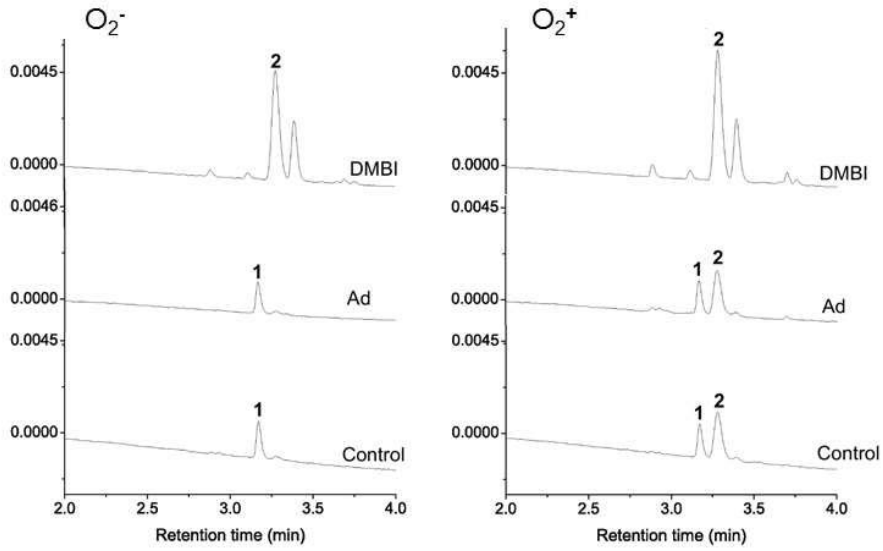

Figure 10 The chromatograms of cobamides obtained from P. freudenreichii cultures grown under strictly anaerobic (O_2^-) and microaerobic (O_2^+) atmospheres, in PPA medium supplemented with only cobalt ions (Control) or additionally adenine (Ad) or DMBI (DMBI). No active vitamin B_{12} (2) was detected in cultures grown under strictly anaerobic atmosphere and only pseudovitamin B_{12} was detected (1) unless DMBI was added. Under microaerobic atmosphere active vitamin B_{12} was formed, however, pseudovitamin B_{12} was still produced, showing that formation of DMBI was a limiting factor. Supplementation with adenine did not increase the amounts of pseudovitamin B_{12} formed compared to control.

5.1.2 RIBOFLAVIN AND NICOTINAMIDE AS SUBSTITUTES FOR DMBI SUPPLEMENTATION (STUDY II)

In study II, 30 strains, 27 *P. freudenreichii* and 3 *A. acidipropionici* were cultivated in a WBM, per the previously tested method aimed at maximising

vitamin B_{12} yields (Hugenschmidt *et al.,* 2010). For this reason, the growth medium was supplemented with 5 mg/L of $CoCl_2$, and the bacteria were cultivated in a two-step fermentation process. In this process bacteria were cultivated anaerobically for the first 72 h, and air was subsequently allowed into the cultures, which was followed by incubation for 96 hours with shaking. Previous studies showed that *P. freudenreichii* synthesises DMBI from riboflavin (Lingens *et al.,* 1992; Renz and Weyhenmeyer, 1972) and that the process is enhanced by supplementation with nicotinamide (Hörig and Renz, 1980). To test whether riboflavin and nicotinamide, the food-grade precursors of DMBI, can substitute for DMBI *in vivo*, cultures supplemented with 40 μM of riboflavin and 27 mM of nicotinamide (Hörig and Renz, 1980) were prepared. The results were compared to the vitamin B_{12} yields from the cultures to which 15 mg/L (100 mM) of DMBI were added. DMBI was added at 144 hours of incubation to prevent inhibition of growth (Hugenschmidt *et al.,* 2010, Marwaha *et al.,* 1983).

5.1.2.1 Bacterial growth in supplemented medium

All strains tested were able to grow in WBM, with and without the supplements (Figure 11A). However, the extent of the growth was strain-dependent. That dependence may be attributed to the strain's ability to utilise

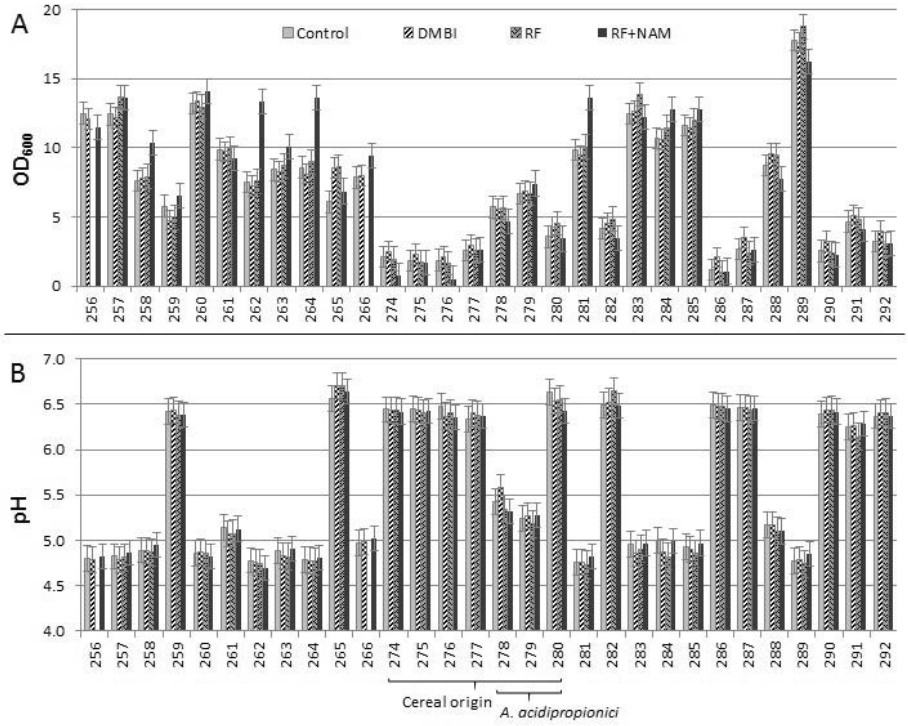

Figure 11 Bacterial growth in WBM (Control) supplemented with DMBI (DMBI), riboflavin (RF) and riboflavin together with nicotinamide (RF+NAM). Panel A: OD600 values recorded after 7 days of incubation; Panel B: pH values recorded after 7 days of incubation. The starting value was 6.5.

lactose, and is reflected by the larger drop of pH from the initial pH 6.5, which resulted from production of the propionic and acetic acids during growth (Figure 11B)(Details in Study II).
Notably, the addition of the precursors did not have a detrimental effect on bacterial growth, with the exception of the riboflavin and nicotinamide combination effect on the growth of two strains, 274 and 276, both of cereal origin. Conversely, the addition of both riboflavin and nicotinamide had a positive effect on the growth of strains 258, 262, 264 and 281.

5.1.2.2 Effect of precursors on production of vitamin B_{12}

All *P. freudenreichii* strains were capable of producing vitamin B_{12} in WBM with cobalt chloride, the control medium. The vitamin B_{12} yield ranged from 0.2 to 5.3 µg/mL, and fell well within the limits reported previously for *P. freudenreichii* (Berry and Bullerman, 1966; Hugenschmidt *et al.,* 2010). The *A. acidipropionici* strains produced the vitamin only when supplemented with DMBI, and even then, the production levels were the lowest of all strains and conditions tested (Figure 12). DMBI supplementation enhanced vitamin B_{12} production in seven of the *P. freudenreichii* strains (256, 257, 281, 283, 288, 289 and 290).

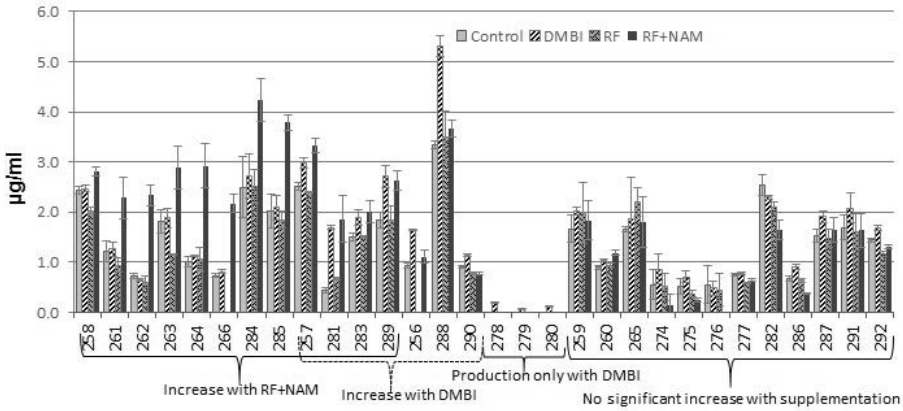

Figure 12 Amount of vitamin B_{12} accumulated in the bacterial cells calculated per mililiter of culture. The strains are grouped according to their response to supplementation.

Addition of riboflavin alone did not increase B_{12} production by the strains, which was attributed to the ~6.6 mM of riboflavin present in the native medium (see Study II). However, it can be assumed that riboflavin contributed to the overall increased yield of vitamin B_{12} because the *P. freudenreichii* strains consumed 25-85% of the riboflavin present in the medium, while the *A. acidipropionici* strains did not(see study II). Supplementation with both riboflavin and nicotinamide resulted in higher production levels of vitamin B_{12} than did supplementation with DMBI for seven strains (261, 262, 263, 264, 266, 284 and 285). Unexpectedly, supplementation resulted in a decrease in vitamin B_{12} production in five strains (274, 275, 276, 282 and 286).

5.1.2.3 Yield of vitamin B_{12} corrected for growth

Production of vitamin B_{12} in *P. freudenreichii* was previously reported as proportional to growth (Hargrove and Leviton, 1955; Vorobjeva, 1999c), which would suggest similar accumulated vitamin levels in the cells. To test this, the obtained amount of vitamin B_{12} was additionally calculated as the amount of vitamin accumulated per gram of wet cells (Figure 13). Thus, the vitamin B_{12} levels accumulated within the cells, and the response to the supplements remained strain dependent. Altogether, these results revealed that the availability of exogenous DMBI was not limiting for B_{12} production under the conditions tested. Additionally, the lack of a response to DMBI supplementation by several strains could be attributed to their inability to take up DMBI from the medium, especially by the strains that responded to the supplementation with riboflavin and nicotinamide. For the other strains, the two-step incubation process itself may ensure sufficient production of DMBI in *P. freudenreichii*. The production capacity appears strain dependent and not directly dependent on growth. *A. acidipropionici* does not appear to produce DMBI and, therefore, must rely on exogenous supplementation to produce active vitamin B_{12}.

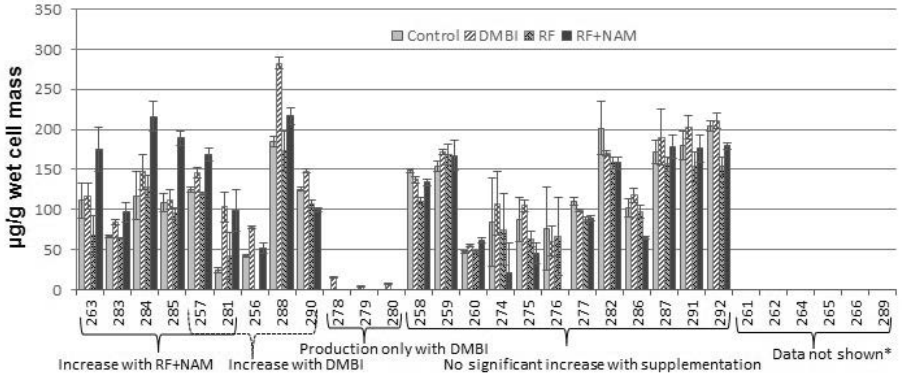

*Figure 13 Amount of vitamin B12 accumulated in the bacterial cells calculated pergram of wet cell mass. The strains are grouped according to their response to supplementation. *No firm cell pellet could be obtained from these strains, and therefore no calculation per gram of cell mass was made.*

5.1.3 VITAMIN B_{12} PRODUCTION UNDER DIFFERENT GROWTH CONDITIONS AND WITHOUT SUPPLEMENTS (STUDY III)

In study III, the capability of *P. freudenreichii* to produce active vitamin B_{12} was assessed in the two-step process, in two different growth media and without cobalt or DMBI addition. The aim was to determine whether the production was affected by the level of growth and the metabolic state of the cells. To this end, the production levels of vitamin B_{12} and other key metabolites were analysed. These analyses were complemented by an analysis of differential global protein production by the *P. freudenreichii 2067 (257)* strain. The samples for the analysis of metabolites were harvested at the mid-

exponential, stationary and late stationary growth phases, while proteomes were analysed from the mid-exponential and stationary phases.

5.1.3.1 Effect of growth media on the level of growth and acid production

We selected two growth media for the study: PPA, a laboratory medium based on sodium lactate and yeast extract, that has been selected for its proximity to nutritional conditions of Swiss cheese (Dalmasso *et al.,* 2012a; 2012b) and WBM, a whey-based medium supplemented with yeast extract, that has often been used for industrial vitamin B_{12} production (Bullerman and Berry, 1966; Marwaha *et al.*, 1983; Hugenschmidt *et al.*, 2011). To determine the growth dynamics in the media, a growth curve that facilitated selection of sampling times for samples corresponding to the mid-exponential phase of growth was created (Figure 14).

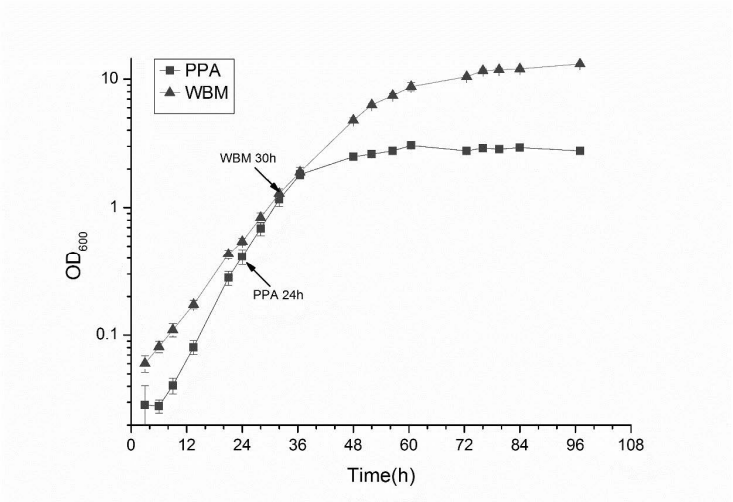

Figure 14 Growth curves used to estimate the time at which cells grown in WBM and PPA reach mid-exponential growth phase. The times selected as mid-exponential phase sampling points for each medium are indicated with arrows.

The growth curve revealed that the media differed markedly in their abilities to support the growth of *P. freudenreichii 2067,* with a considerably shorter generation time of 4.2 hours in PPA versus 5.8 hours in WBM, but with a much higher OD_{600} value of 10.3 in WBM versus 2.8 in PPA at the end of growth. Additionally, based on the growth curve, sampling points at the mid-exponential phase were established at 24 hours for PPA and 30 hours for WBM.

The differences in the amounts of available carbon sources in the two media were reflected by the acid production (Figure 15). The abundance of lactose resulted in much higher concentrations of propionic acid (13.8 ± 1.3) and acetic acid (5.4 ± 0.1 mg/mL) in WBM, which was concurrent with the pH drop from 6.4 to 4.6. In agreement with previous studies (Piveteau, 1999; Crow, 1986), in WBM, *P. freudenreichii 2067* fermented lactic acid prior to

fermenting lactose. The presence of unused lactose (25 ± 1.5 mg/mL) in WBM at the end of incubation suggests that in this medium growth was inhibited by low pH rather than by exhaustion of carbon sources, as opposed to PPA, where pH increased from 6.0 to 6.2. In PPA, the starvation imposed by the limited amount of carbon sources-first lactic acid and then pyruvic acid- prompted swith to amino acid metabolism in *P. freudenreichii 2067*, as evidenced by the increased amount of acetic acid (2.9 ± 0.0 mg/mL) in proportion to the propionic acid (5.9 ± 0.2 mg/mL) accumulated (Crow, 1986).

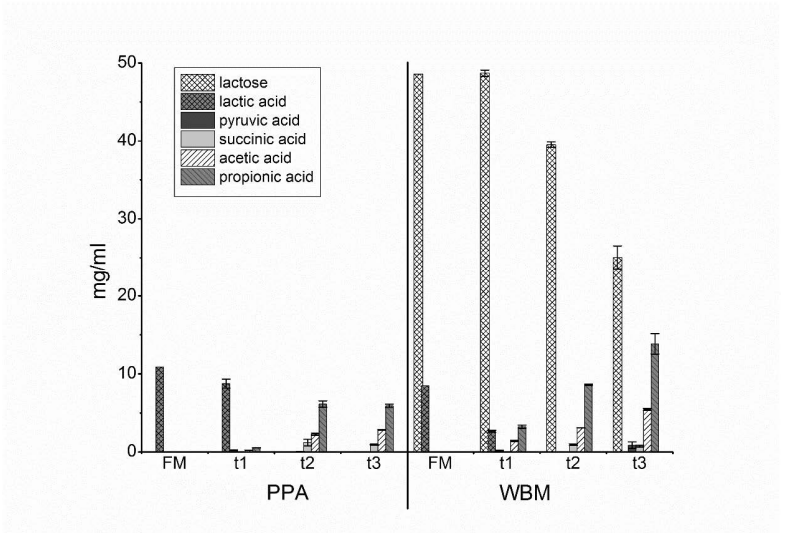

Figure 15 Levels of sugars and accids accumulated in the PPA and WBM growth media during growth of P. freudenreichii 2067 at mid-exponential growth phase (t1), stationary phase (t2) and at the end of incubation at 168 hours (t3). The levels detected in fresh media are also shown (FM).

5.1.3.2 Production of active and pseudovitamin B_{12}

The levels of active and pseudovitamin B_{12} measured from the *P. freudenreichii* cells at the mid-exponential growth phase were similar in both media, at <20 ng/mL and <7 ng/mL, respectively (Figure 16).

At 72 hours, the cultures were transferred to aerobic growth, and concurrently stationary phase samples were taken. At this time point, slightly higher levels of vitamin B_{12} (45.3 ± 1.7 ng/mL) were detected in PPA relative to WBM (37.9 ± 1.9 ng/mL), while the levels of pseudovitamin B_{12} remained similar in both cultures (below 40 ng/mL). At the late stationary phase, levels of vitamin B_{12} detected in WBM and PPA were 124.8 ± 34.7 ng/mL and 120 ± 10.4 ng/mL, respectively. No pseudovitamin B_{12} was detected at this time point. It was assumed that the increased availability of oxygen allowed for the production of endogenous DMBI and led to apparent conversion of pseudovitamin B_{12} into the active form. The production of active vitamin during anaerobic incubation was attributed to the availability of residual oxygen in the cultures.

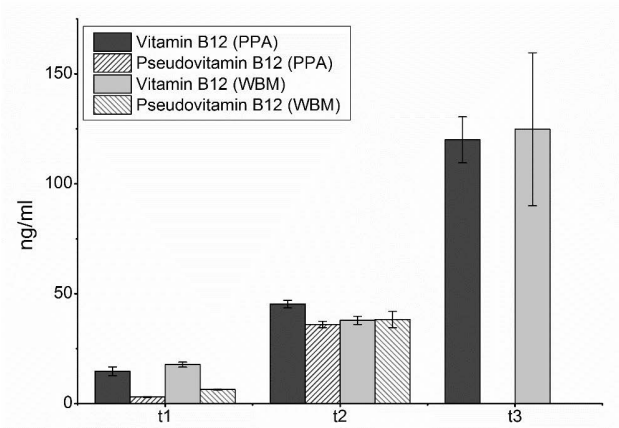

Figure 16 Production of cobamides measured from the cells of P. freudenreichii 2067 grown in PPA and WBM media. The sampling times were mid-exponential growth phase (t1), stationary phase 72h (t2) and at the end of incubation at 168 hours (t3). No pseudovitamin B_{12} was observed after the switch to aeration.

Although the anaerobic jars ensure the removal of oxygen from the headspace of the culture, the oxygen that is dissolved in the medium remains.

For the relative comparison, the amount of vitamin B_{12} produced in the cultures was recalculated as the amount of vitamin accumulated per gram of wet cells (Figure 17). The production of vitamin was found to be higher in PPA-grown cells than in WBM-grown cells at each sampling point, suggesting that at the growth conditions tested, the vitamin B_{12} production is not dependent on growth. Conversely, the production is most likely limited by the restricted availability of cobalt, as the strain 2067 (257, see Table 3) grown under the same conditions, but with addition of cobalt, accumulated over 10-fold more of vitamin B_{12} per gram of cell mass (see figure 13).

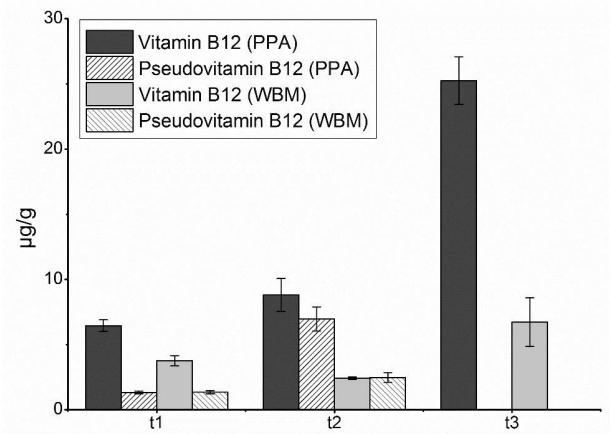

Figure 17 Production of cobamides accumulated per gram of wet cell mass. The sampling times were mid-exponential growth phase (t1), stationary phase 72hours (t2) and at the end of incubation at 168 hours (t3). The lower growth in PPA resulted in higher amounts of vitamin B_{12} accumulated per gram of cell mass.

5.1.3.3 Global protein production

The 2DE analysis of the global protein production during the mid-exponential phase in PPA- and WBM-grown *P. freudenreichii* 2067 cells led to the identification of 51 gene products that were identified from 28 differentially abundant protein spots (fold change ≥ 1.4, $p < 0.05$) (Table 5 and study III, Supplemental file 1). Among these, 12 proteins were identified as single hits.

Eight of the proteins were more abundant in the PPA proteome; they included proteins associated with lipid metabolism (FabH, MaoC), energy production (AtpD, FixB), and protein production (GlnRS and MiaB1). Two potential oxidoreductases (the SDR and FAD-dependent pyridine nucleotide-disulfide oxidoreductases) and a putative glutamate synthase (GLS) were also identified. Proteins identified as more abundant from the WBM proteome included oxidative stress-associated cysteine synthase (Cys2) and a superoxide dismutase (SodA), each identified from two separate spots.

The analysis of the protein production during the stationary phase returned 112 gene products from 65 differentially abundant spots (fold change ≥ 1.5, $p < 0.05$), with 21 single identifications (Table 5). The more abundant proteins from the PPA proteome included proteins involved in energy generation through the metabolism of pyruvate (NifJ1, PPDK) or feeding of the TCA cycle (SdhA3, AspA1/2, and GLS). Also, the stress response-related proteins ClpATPase (ClpB2), protein synthesis-associated SerRS1 and a potential dehydrogenase (SDR_c) were found. The proteins identified from the WBM proteome included proteins associated with glycolysis/lactate utilisation (L-LDH, GAPDH, FBA2), nitrogen metabolism (GDH), and stress (SDR, ClpB1, Cys2, SodA, IbpA).

The differences in the proteomes reflected the differences in the media. For example, the cells grown in PPA without lipid sources had elevated levels of proteins involved in lipid metabolism, while the faster growth rate was reflected in the higher abundance of enzymes involved in protein synthesis. However, the abundance of lactose and resulting high acid production led to increased production of several stress-related proteins (ClpB1, Cys2, and SodA) in WBM. The switch to lactose utilisation at the later growth stages was reflected by the increased abundance of glycolytic proteins such as GAPDH and FBA2, while the increased production of two aspartate ammonia lyases during the stationary phase in PPA suggested a switch to amino acid metabolism for energy generation in response to the carbon source limitation.

Out of over 20 vitamin B_{12} biosynthetic proteins, five (HemB, HemL1, HemL2, CbiL and CbiO) were identified from spots that were differentially abundant in the PPA versus WBM proteomes (see study III, Supplemental file 1). However, neither of these proteins were detected as a single or best hit. Therefore, it can be assumed that under the conditions tested, the abundance of B_{12} biosynthesis proteins did not significantly differ between PPA- and WBM-grown cells.

Table 5 List of proteins detected by LC-MS/MS from spots with relative abundance fold change ≥1.4 and ≥1.5 (ANOVA, $p \leq 0.05$) from PPA and WBM media during exponential and stationary growth phases, respectively. Only spots with single identification are presented

Spot no.	Fold change	Mw	pI	Mascot score	No. of match. pept.	Seq. cov.	Locus in ref. organism	Name	Annotation
					Exponential phase				
	PPA								
2	1.5	64546	5.2	709	9	22%	PFREUD_11940	GlnRS	Glutaminyl-trna synthetase
5	1.4	60757	5.5	1691	33	65%	PFREUD_01830	GLS	Putative glutamate synthase
6	1.5	56027	5.8	274	2	20%	PFREUD_13090	MiaB1	2-methylthioadenine synthetase
8	2.2	52579	4.7	1706	30	69%	PFREUD_10490	AtpD	ATP synthase subunit beta
14	1.5	31894	4.8	487	9	15%	PFREUD_09430	FabH	3-oxoacyl-ACP synthase
15	1.4	32914	5.4	1038	27	65%	PFREUD02450	MaoC	Maoc acyl dehydratase
16	2.3	29633	4.5	676	12	51%	PFREUD_02490	FixB	Electron transfer flavoprotein fixb
18	1.6	23335	5	530	7	44%	PFREUD_23970	SDR	Oxidoreductase
	WBM								
24	1.5	33620	5.1	903	16	58%	PFREUD_16420	Cys2	Cysteine synthase A
25	1.9	33620	5.1	1557	34	82%	PFREUD_16420	Cys2	Cysteine synthase A
26	1.5	22619	5.3	392	4	26%	PFREUD_06110	SodA	Superoxide dismutase
27	1.5	22619	5.3	611	16	47%	PFREUD_06110	SodA	Superoxide dismutase
					Stationary phase				
	PPA								
29	1.6	1E+05	5.2	2689	40	33%	PFREUD_01840	NifJ1	Pyruvate synthase/pyruvate-flavodoxin oxidoreductase
30	1.9	1E+05	5.2	2525	59	34%	PFREUD_01840	NifJ1	Pyruvate synthase/pyruvate-flavodoxin oxidoreductase
31	1.7	1E+05	5.2	3124	71	42%	PFREUD_01840	NifJ1	Pyruvate synthase/pyruvate-flavodoxin oxidoreductase
36	1.7	96275	4.8	2351	50	39%	PFREUD_03230	PPDK	Pyruvate phosphate dikinase
37	2.3	94317	5.3	2756	48	44%	PFREUD_17920	ClpB2	ATP-dependent Clp protease B2

Spot no.	Fold change	Mw	pI	Mascot score	No. of match. pept.	Seq. cov.	Locus in ref. organism	Name	Annotation
39	2.4	75276	6	1488	36	36%	PFREUD_14311	SdhA3	Succinate dehydrogenase flavoprotein subunit
42	1.9	60757	5.5	1966	67	62%	PFREUD_01830	GLS	FAD-dependent pyridine nucleotide-disulfide oxidoreductase
43	3.6	53385	5.1	825	62	25%	PFREUD_16320	AspA1	Aspartate ammonia-lyase
43	3.6	53526	5.1	748	48	26%	PFREUD_16330	AspA2	Aspartate ammonia-lyase
46	2	47518	5	270	4	11%	PFREUD_01350	SerRS1	Seryl-trna synthetase
47	1.5	47953	5.7	678	11	27%	PFREUD_05140	GGR	Electron transfer oxidoreductase (Geranylgeranyl reductase superfamily)
57	1.6	24018	5.2	352	21	33%	PFREUD_22440	SDR_c	Short subunit dehydrogenase (classical)
	WBM								
59	2	93458	5.1	2766	46	50%	PFREUD_19250	ClpB1	ATP-dependent Clp protease B1
60	1.5	93458	5.1	2527	49	46%	PFREUD_19250	ClpB1	ATP-dependent Clp protease B1
70	2	49226	5.5	1290	20	51%	PFREUD_17610	GDH	Glutamate dehydrogenase
75	3	34225	4.9	661	15	31%	PFREUD_12840	L-LDH	L-lactate dehydrogenase
78	1.6	35532	5.2	609	8	27%	PFREUD_15130	GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
81	2	32368	5	582	8	34%	PFREUD_23890	FBA2	Fructose-bisphosphate aldolase class I
85	3.7	33620	5.1	431	6	28%	PFREUD_16420	Cys2	Cysteine synthase A
86	2.2	33620	5.1	826	47	54%	PFREUD_16420	Cys2	Cysteine synthase A
91	1.8	23335	5	476	6	46%	PFREUD_23970	SDR	Oxidoreductase
92	1.5	22619	5.3	314	6	18%	PFREUD_06110	SodA	Superoxide dismutase
93	5.2	16830	4.9	575	8	58%	PFREUD_09500	IbpA	Molecular chaperone (small heat shock protein)

Taken together, these results imply that *P. freudenreichii* is capable of producing active vitamin B_{12} regardless of the metabolic state of the cells and that the production is limited by the availability of cobalt. Nevertheless, the levels of vitamin B_{12} obtained are relevant from the perspective of food fortification. Absorption of vitamin B_{12} from food is mediated by the Intrinsic Factor, and it reaches a maximum at 1.5-2 μg per meal (EFSA NDA Panel, 2015). With the absorption level of vitamin B_{12} estimated at 40% from food (EFSA NDA Panel, 2015), the content of vitamin B_{12} of 4-5 μg per food portion would be sufficient and is in line with the adequate intake (AI) value of 4-5 μg/day for adults proposed by the EFSA. These levels could be reached with 34-42 mL of the cultures with the vitamin B_{12} production levels obtained here.

5.2 GENOMICS OF *P. FREUDENREICHII*

At the beginning of this study only one complete genome sequence was available for *P. freudenreichii* (Falentin *et al.,* 2010). The most likely reason for this was the lack of a suitable sequencing platform. Propionibacteria have DNA with high GC% contents and multiple regions of repeated sequences. These two features mean that PCR-based sequencing is unreliable due to the so-called GC bias, while the platforms that produce short sequence reads render the genome assembly difficult, if possible at all (Lee *et al.,* 2016). For this reason, in study IV, the PacBio RSII sequencing platform, in which no GC bias is observed and in which the long reads allow genome closing (Rhoads and Au, 2015) was tested on *P. freudenreichii* DSM 20271. In study V, additional 18 strains were sequenced and a comparative genomics study was performed on the 20 complete genomes of *P. freudenreichii* that are currently available.

5.2.1 WHOLE GENOME SEQUENCING WITH PACBIO (STUDY IV)

In study IV, whole genome sequencing of one of the *P. freudenreichii* type strains, *P. freudenreichii* subsp. *freudenreichii* DSM 20271 was performed. The sequence was generated on the PacBio RSII platform using a standard PacBio 10 kb library, two SMRT cells and a 180 min movie time. With the HGAP algorithm (Chin *et al.,* 2013), the genome was assembled into a single, circular contig (Figure 18), confirming that the PacBio sequencing platform

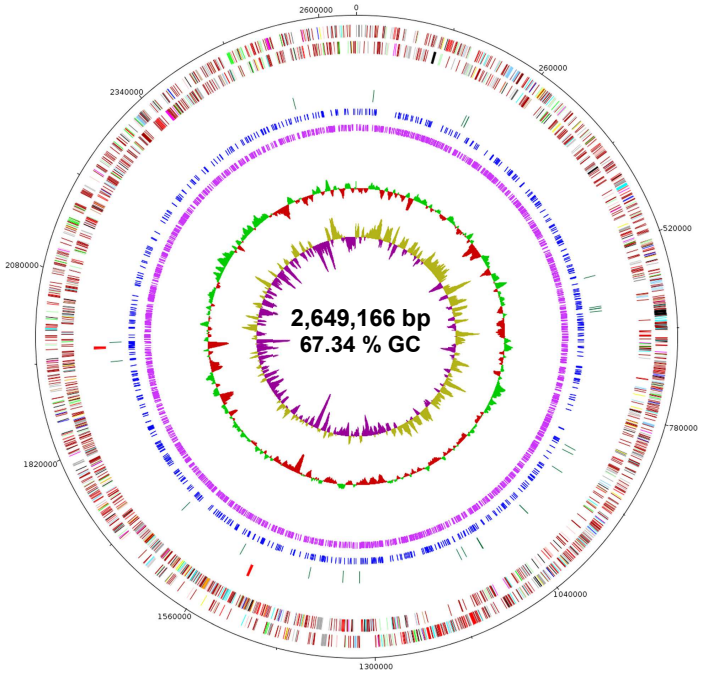

Figure 18 Genome properties of Propionibacterium freudenreichii ssp. freudenreichii DSM 20271T . The circles are from inner to outer order: GC skew, GC%, putative m4C, m6A, tRNA, rRNA, CDS reverse, CDS forward. CDS are colored according to the COG functional categories

can overcome the high GC% content and the regions of repeated sequences in *P. freudenreichii* and showing that it is well-suited for *de novo* whole genome sequencing of the species. Following the publication of study IV, an additional whole genome sequence of the strain *P. freudenreichii JS* closed with the aid of the PacBio sequencing was published (Ojala *et al.,* 2017).

5.2.2 SEQUENCING OF 18 ADDITIONAL *P. FREUDENREICHII* STRAINS (STUDY V)

In study V, 17 new genomes were sequenced and the type strain DSM 4902 (JS15) was resequenced. Seven of the strains were sequenced with the use of P4/C2 chemistry, while for the remaining 11 the newer P5/C3 chemistry was used (Table 6).

Table 6 Summary of the sequencing parameters used to obtain the whole genome sequences of the 18 P. freudenreichii strains.

Strain name	Sequencing Chemistry	SMRT cells	Movie Time (min)	Total Number of Subreads	Total Number of Bases	Mean Subread Length (bp)	N50 Subread Length (bp)
JS4	P4/C2	2	120	142,306	465,129,578	3,268	4,131
JS10	P4/C2	2	120	205,113	606,802,054	2,958	3,733
JS15	P4/C2	2	120	179,642	540,206,310	3,007	3,653
JS18	P4/C2	2	120	194,246	550,199,736	2,832	3,679
JS20	P4/C2	2	120	230,371	632,635,668	2,746	3,456
JS22	P4/C2	2	120	181,361	563,541,807	4,090	3,107
JS23	P4/C2	2	120	192,373	617,310,336	3208	4,222
JS2	P5/C3	2	240	100,864	719,516,424	7,133	8,940
JS7	P5/C3	2	240	107,194	1,002,978,417	9,356	13,043
JS8	P5/C3	2	240	139,418	1,149,552,255	8,245	11,148
JS9	P5/C3	2	240	160,808	1,316,084,462	8,184	11,023
JS11	P5/C3	2	240	184,744	1,422,323,823	7,698	10,498
JS12	P5/C3	2	240	207,391	1,490,421,212	7,186	9,765
JS13	P5/C3	2	240	103,466	793,823,369	7,672	10,156
JS14	P5/C3	2	240	125,130	966,521,587	7,724	10,326
JS17	P5/C3	2	240	151,002	1,149,295,071	7,611	10,232
JS21	P5/C3	2	240	173,843	1,298,367,718	7,468	9,906
JS25	P5/C3	2	240	192,104	1,293,994,542	6,735	8,805

The newer chemistry allowed for an extended sequencing movie time, increased mean read length and a higher total of sequenced bases. Overall, it resulted in a deeper genome coverage (Table 7). Nevertheless, both chemistries allowed for the assembly of the genomes into single, circular contigs.

Assembly of the sequencing data obtained from the 18 strains resulted in 31 circular sequences. Eight of the strains had two complete chromosomes, which stemmed from duplication and relocation of transposable elements (Table 8).

Table 7 The P. freudenreichii strains included in study V. Table shows the summary of the 31 genomes sequenced in this study from 18 strains as well as the genomes of the 2 previously sequenced strains JS and DSM 20271. In case of the strains with two bacterial genomes, the sequences used for the comparative genomics are marked with asterisk.

Strain name	Sequence name	Accession number	Genome size (bp)	Genome coverage	GC%	No. of predicted genes	Note
JS	PFREUDJS001	LN997841.1	2 675 045	691	67%%	2 382	Sequenced previously (Ojala *et al.* 2017)
JS2	PFRJS2	LT576032	2 655 351	317	67%	2 310	
JS4	PFRJS4	LT576033	2 654 663	158	67%	2 366	
JS7	PFRJS7-1*	LT618776	2 738 418	300	67%	2 412	
	PFRJS7-2	LT618777	2 700 482	50	67%	2 355	
	PFRJS7-PH	LT618778*	37 936	60	65%	59	Phage genome
JS8	PFRJS8	LT576042	2 655 373	392	67%	2 324	
JS9	PFRJS9-1*	LT618785	2 720 049	219	67%	2 405	
	PFRJS9-2	LT618786	2 718 592	219	67%	2 401	
JS10	PFRJS10	LT576035	2 626 110	208	67%	2 329	
JS11	PFRJS11	LT576038	2 537 402	501	67%	2 200	
JS12	PFRJS12-1*	LT604998	2 615 181	250	67%	2 275	
	PFRJS12-2	LT576787	2 613 734	250	67%	2 274	
	PFRJS12-3	LT604882**	24 909	400	64%	32	Putative conjugtive plasmid
JS13	PFRJS13-1*	LT618779	2 537 370	210	67%	2 201	
	PFRJS13-2	LT618780	2 520 651	90	67%	2 189	
JS14	PFRJS14	LT593929	2 507 188	330	67%	2 180	
JS15	PFRJS15-1*	LT618787	2 621 081	94	67%	2 322	
	PFRJS15-2	LT618788	2 616 005	94	67%	2 320	
JS16	RM25	NZ_CP010341.1	2 649 163	190	67%	2 321	Sequenced previously (Koskinen *et al.* 2015)
JS17	PFRJS17-1*	LT618789	2 755 516	192	67%	2 455	
	PFRJS17-2	LT618790	2 754 069	192	67%	2 454	
JS18	PFRJS18	LT576034	2 661 974	190	67%	2 358	
JS20	PFRJS20-1*	LT618791	2 678 207	106	67%	2 376	
	PFRJS20-2	LT618792	2 682 327	106	67%	2384	
JS21	PFRJS21-1*	LT618781	2 659 993	222	67%	2 330	
	PFRJS21-2	LT618782	2 658 550	222	67%	2 329	
JS22	PFRJS22-1	LT599498	2 633 661	190	67%	2 326	
	PFRJS22-PH	LT615138*	39 309	102	66%	61	Phage genome
JS23	PFRJS-23-PH	LT618794	42 723	25	65%	66	
	PFRJS-23	LT618793*	2 630 698	210	67%	2 335	Phage genome
JS25	PFRJS25-1	LT618783	2 666 517	400	67%	2 336	
	PFRJS25-2	LT618784**	35 640	800	64%	46	Putative conjugative plasmid

Interestingly, the transposases did not appear to be acquired through horizontal gene transfer, as each of them was present in their respective genomes in at least one additional location. Moreover, all of these transposases had counterparts in multiple strains, suggesting their intrinsic character in the species and implying a role in the adaptation strategies.

Table 8 Details of the differences found between the genome sequences in the strains for which more than one bacterial genome sequence was obtained.

Strain name	Accession number	Note	Details
JS7	LT618776	Genome with prophage	PFR_JS7-1_1810 HTH-type transcriptional regulator KmtR: PFR_JS7-1_1869 Transcriptional regulator MtrR (preceeded by tRNA-Lys)
	LT618777	Genome without phage	
JS9	LT618785	Genome with an additional transposase	PFR_JS9-1_62 Uma4 protein
	LT618786		
JS12	LT604998	Genome with an additional transposase	PFR_JS12-1_615* Transposase of ISAar20, ISL3 family and PFR_JS12-1_616 Hypothetical protein (ahead of Carbon starvation protein)
	LT576787		
JS13	LT618779	Genome with 12-gene insertion or deletion**	JS13_289 Hypothetical protein: JS13_299 Transposase for insertion sequence element IS1001
	LT618780		
JS15	LT618787	3 transposases absent from the other genome	PFR_JS15-1_878 Uma4 protein and PFR_JS15-1_879 Hypothetical protein; PFR_JS15-1_1737 Transposase of ISAar20, ISL3 family (Transposase disrupting gene coding for CitT); PFR_JS15-1_2045* Transposase of ISAar20, ISL3 family
	LT618788	Transposase disrupting a Type III RM methylase	PFR_JS15-2_359 Transposase of ISAar43, IS3 family, IS407 group, orfA and PFR_JS15-2_360*** Insertion sequence IS407 orfB
JS17	LT618789	A duplicated transposase	PFR_JS17-1_657* Transposase of ISAar20, ISL3 family
	LT618790		
JS20	LT618792	3 transposases absent from the other genome	PFR_JS20-2_544 Insertion sequence IS407 OrfB and PFR_JS20-2_545*** Transposase of ISAar43, IS3 family, IS407 group, orfA (disrupting Transposase IS30 family); PFR_JS20-2_568 Transposase of ISAar20, ISL3 family (ahead of Carbon starvation protein); PFR_JS20-2_1234 Transposase of ISAar20, ISL3 family (ahead of LuxS)
	LT618791		
JS21	LT618781	Additional transposase	PFR_JS21-2_248 Transposase of ISAar20, ISL3 family
	LT618782		

*Transposases PFR_JS12-1_615, PFR_JS15-1_2045 and PFR_JS17-1_657 found in strains JS12, JS15 and JS17, respectively, are 100% identical.
**The same element is present in the strain JS11 in the same location, but in all of the sequences.
*** Transposases PFR_JS15-2_359 and PFR_JS15-2_360 are identical to transposases PFR_JS20-2_544 and PFR_JS20-2_545

The two genomes of the strain JS13 differed by 12 genes that were organised into a gene cluster. Aside from 4 transposases, the gene cluster consisted of eight hypothetical proteins, one with similarity to Helicase conserved C-terminal domain (PF00271.25). It was unclear whether the presence/absence of the cluster resulted from an insertion or deletion, but an identical gene cluster was found in one additional strain, JS11, in 100% of the genome sequences.

Three of the sequences were presumably mobilised prophages, as their chromosomal counterparts were also found. The strain JS7 was the only phage-carrying strain that also had a genome from which the phage was absent. Two additional sequences were classified as putative conjugative plasmids, carried by the strains JS12 and JS25.

A large duplication was detected in the genome of the JS17 strain. The duplication appeared transposase-mediated and spanned 35 genes: PFR_JS17-1_676 to PFR_JS17-1_710 and PFR_JS17-1_711 to PFR_JS17-1_745, located between an Uma4-type transposase and an aspartate ammonia lyase. In order to eliminate the possibility of an assembly error, the edges of the duplication by PCR were analysed. The duplication region included genes coding for, among others, thiamine biosynthetic proteins, transporters and glycerol metabolism proteins. Most notably, the duplication led to the presence of three copies of aspartate ammonia lyase instead of the usual two copies for *P. freudenreichii*, which could result in increased aspartase activity in the JS17 strain. The activity of aspartate ammonia lyase is associated with gas production and formation of desirable eyes in cheeses. However, it can lead to excessive gas production and result in a split defect (Daly *et al.,* 2010).

5.2.2.1 Evaluation of re-sequencing of the strain DSM 4902

The suitability of the sequencing method can be assessed by comparing the sequencing outcome with a sequence obtained through a high-confidence method such as Roche 454 or Sanger (Chin *et al.,* 2013; Koren *et al.,* 2013). Therefore, the complete genome sequence for the strain ATCC9614 (FN806773.1), which was obtained through Sanger sequencing (ABI3730 sequencer) and assembled with primer walking (Falentin *et al.,* 2010) was compared with the PacBio-sequenced version, strain Js15 (LT6187879) (Figure 19). The BLASTn search revealed 99% identity over 99% of the

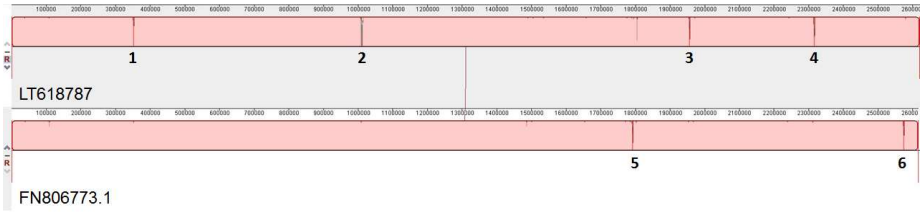

Figure 19 Progressive Mauve alignment of the two genome sequences of the strain P. freudenreichii DSM 4902, the one generated with PacBio platform (LT6187879) and with Sanger sequencing (FN806773.1). The numbered features correspond to regions of sequences missing from the other version of the genome. The resulting "gaps" are listed in Table 9.

sequence. The Progressive Mauve alignment of the genomes allowed an analysis of the differences between the genomes by revealing the gaps and single-nucleotide polymorphisms (SNPs).

Table 9 Summary of the gaps between the two versions of the genome sequences of the strain P. freudenreichii DSM 4902.

Feature on the alignment	Genome	Position in the genome	Gap length	Position in the other genome	Locus ID in the other genome	Annotation
1	FN806773.1	351844	820	351604	PFR_JS15-1_292	Transposase for insertion sequence element
2	FN806773.1	1008180	3195	1008757	PFR_JS15-1_878	Uma4 protein
3	FN806773.1	1953026	1433	1955599	PFR_JS15-1_1737	Transposase of ISAar20, ISL3 family
4	FN806773.1	2311186	1707	2315190	PFR_JS15-1_2045	Transposase of ISAar20, ISL3 family
5	LT618787	1795117	1254	1791345	PFREUD_24740, PFREUD_24750	2 transposases, repeat region
6	LT618787	2579351	1255	2573640	PFREUD_25150, PFREUD_25160	2 transposases, repeat region

The analysis suggests that the major differences lay in the presence and locations of different transposases in each of the genomes (Table 9), while the minor differences, not visible on the alignment, were attributable to the nine identified SNPs (Table 10). The observed differences can be a result of the origins of the strains, which although equivalent, were obtained from different strain collections. This comparison confirms that both sequencing methods are suitable for whole genome sequencing of *P. freudenreichii*.

Table 10 Summary of the SNPs between the two genome sequences of the strain DSM 4902

SNP pattern	Position in the genome LT618787	Position in the genome FN806773.1
cg	37260	37500
ca	107572	107812
ct	1488990	1485221
ga	1503537	1499768
ac	1656803	1653033
ga	1774109	1770339
ga	1788218	1784448
ct	1970134	1966130
cg	2242225	2238221

5.2.3 COMPARATIVE GENOMICS

The newly sequenced genomes and the two previously published sequences for strains JS16 (study IV, Koskinen *et al.,* 2015) and JS (Ojala *et al.,* 2017) were subjected to a comparative genomics analysis.

5.2.3.1 Genome organization and average nucleotide identity

The ANI calculated through pairwise BLAST alignments (Richer and Rosselló-Móra 2009) revealed that the genomes of the *P. freudenreichii* strains were highly collinear, with an ANI value of 97.96% between strains JS7 and JS9 and 100% between strains JS11 and JS13 (Figure 20A). Based on the ANI values, strains JS4, JS15 and JS17 were on average the most similar to all

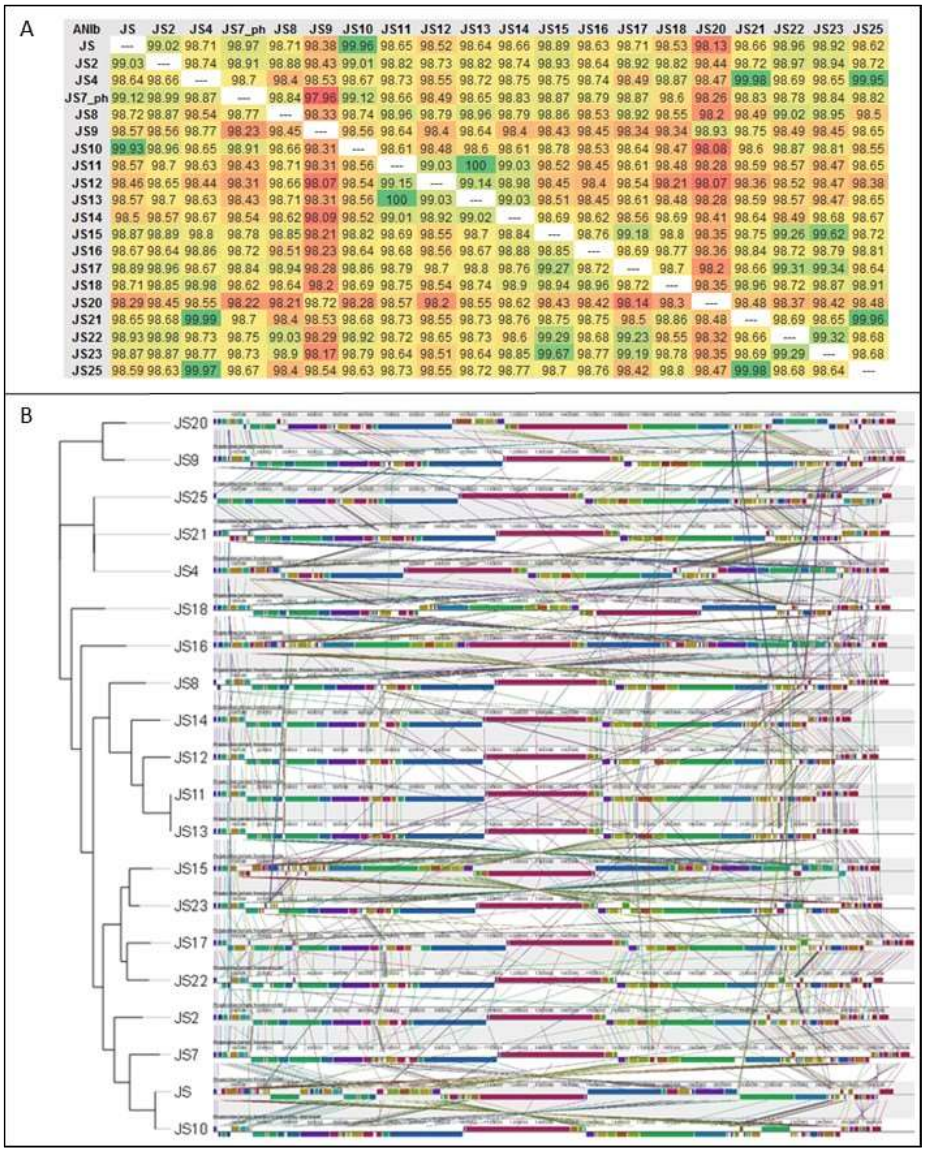

Figure 20. Genome composition and organisation. Panel A) Average Nucleotide Identity (%) calculated based on pairwise BLAST alignment (ANIb). The levels of similarity are highlighted by coloring from green for the most similar to red for the most dissimilar; Panel B) Whole genome alignments generated with ProgressiveMauve. The genomes are arranged according to the phylogenetic tree generated from core genome alignments (see "Core and pangenome analyses").

other strains, while strains JS9 and JS20 were the most dissimilar to all other strains, while they were slightly more similar to one another.

The strains of cereal origin (JS11-JS14) were more similar to each other than to other strains. The whole genome alignments with Progressive Mauve showed that despite genome-wide collinearity, large regions of inversions and other types of reorganisations were present, even among closely related strains. Such reorganisations were observed most clearly between strains JS and JS10, JS15 and JS23, and JS4 and JS21 (Figure 20B). These genome reorganisations appear to be transposase-mediated, as transposases are found on the edges of several of the locally collinear blocks. It is noteworthy that both strain types (JS15 and JS16) show a markedly different genome organisation from the other strains, possibly because of adaptation to the extended exposure to laboratory conditions.

5.2.3.2 Core genome and pangenome analyses

For the estimation of the core genome and pan-genome, one genome from each strain was re-annotated with PROKKA and analysed with the Roary pipeline (Page *et al.,* 2015). The analysis revealed 4606 orthologue groups out of which 1636 belonged to the core genome. The number of orthologue groups in accessory genomes varied between 530 in strain JS12 to 787 in strain JS17. The strain JS20 had the highest number of unigue genes (160), followed by strains JS7 (135), JS12 (129), JS17 (118) and JS9 (116). The strains JS11 and JS13 have one and no unique genes, respectively, due to the fact that the strains are largely identical. In addition to the 12 genes missing from part of the genome sequences of the strain JS13, the differences between those strains include: the gene PFR_JS11_1775 coding for Peptidase dimerization domain protein is split into two in strain JS13 (PFR_JS13-1_1775 and PFR_JS13-1_1776), the gene PFR_JS13-1_1785 coding for Hypothetical protein is missing from the JS11 strain, while the gene PFR_JS11_2200 coding for Hypothetical protein is missing from the strain JS13. The numbers of orthologues belonging to the accessory genome of each strain as well as unique genes are shown in Figure 21.

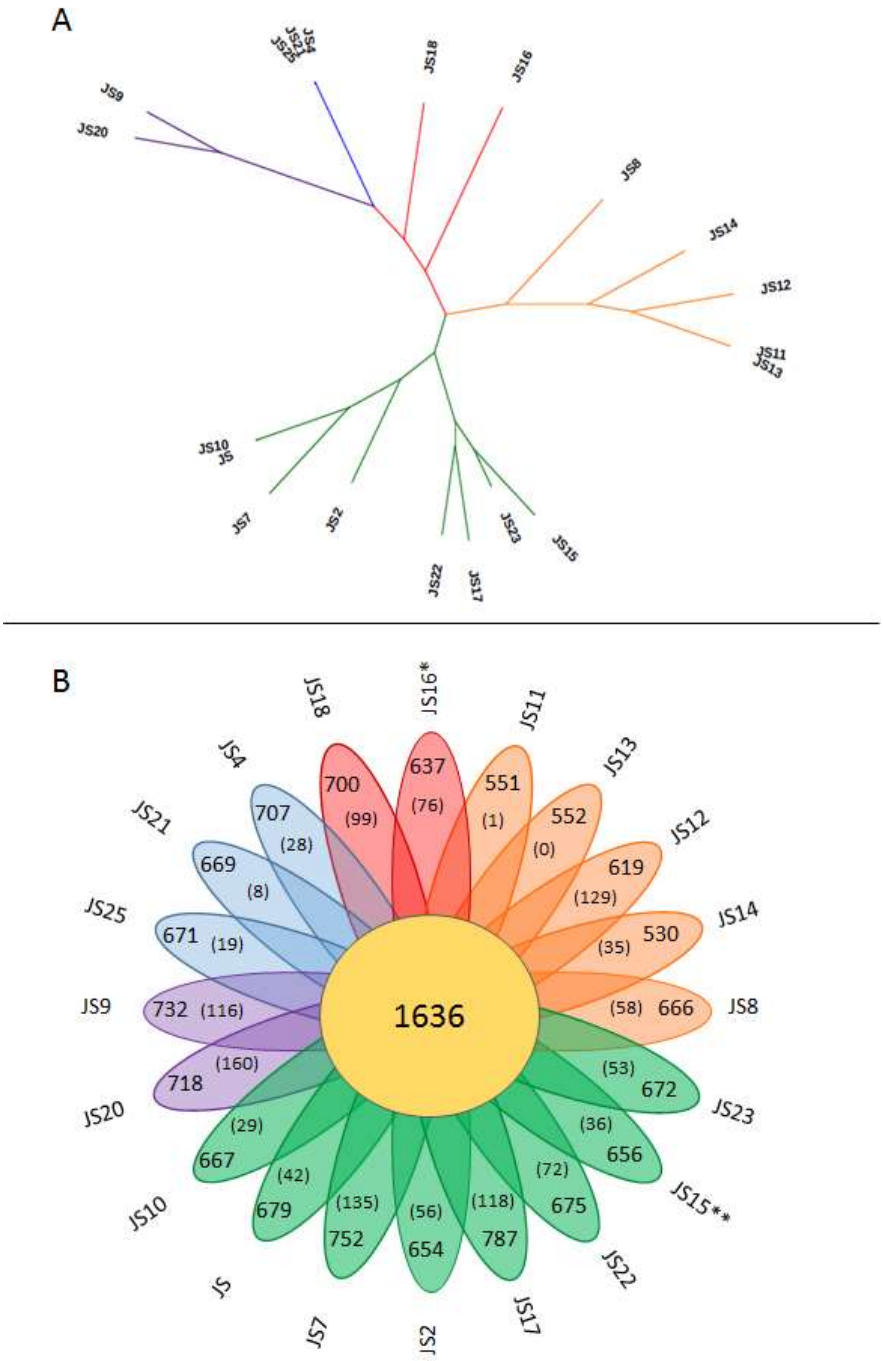

Figure 21 Panel A) An unrooted phylogenetic tree constructed on the basis of core genome alignments; Panel B) Flower plot representing comparative analysis of the genomes. The number of orthologues in the accessory genome of each strain are indicated on the petals. The number of orthologues unique to the strain is indicated in brackets. The petals are coloured based on the phylogenetic grouping of the strains.

**Type strain P. freudenreichii DSM 20271; **Type strain P. freudenreichii DSM 4902*

5.2.4 BIOINFORMATICS ANALYSES

The genomes were analysed further with multiple bioinformatics tools, allowing for more in-depth characterisation of individual strains. This approach was used to identify the mobile elements, including bacteriophages, plasmids and genomic islands. Then, the strains' immunity to mobile elements, namely the CRISPR-Cas systems as well as restriction-modification (RM) systems were explored.

5.2.4.1 Bacteriophages

Relatively little is known about bacteriophages that infect *P. freudenreichii*. Currently, only ten complete bacteriophage genome sequences are available: one filamentous phage (B5), and nine tailed phages (B3, B22, E1, E6, G4, Doucette and Anatole (Modlin *et al.*, unpublished) and PFR1 and PFR2 (Brown *et al.*, 2016). In the current study, three additional bacteriophage genomes were detected as a circular DNA within strains JS7, JS22 and JS23 (PJS7, PJS22 and PJS23) (see Table 7). The phage found in the strain JS7 (LT618778) had the total genome size of 37936 bp and 59 predicted open reading frames. When integrated into the chromosome as a prophage, it was located between the coding sequence for transcriptional regulator KmtR, located immediately downstream of a tRNA-Ala(agc) and a tRNA-Lys(ttt), located immediately upstream of the transcriptional regulator MtrR. A BLAST search against the known Propionibacteria phages showed that PJS7 was 99% identical with the 38071-bp-long PFR1 phage (NC_031076.1). The difference could be found in the gene that coded for the minor tail protein, where PFR-JS7_47 was 135 nucleotides shorter than the gene coding for BI042_gp13 of the PFR1 phage.

Strain JS7 differed from the other two phage-carrying strains by the additional presence of a bacterial genome that was cleared of a prophage (LT618777). Replication of the circular phage genome in JS7 was followed by PCR after subculturing, which revealed the successive integration of the phage genome at the same locus after five passages in the PPA medium (Table 11).

The other two strains, JS22 and JS23 carried the prophage on all of the copies of the chromosome. Phages PJS22 and PJS23 were 97% identical over 68% of their sequences. PJS22 showed 99% identity over 81% of the sequence to previously sequenced phage B22 (KX620750.1). PJS23 was the most similar to phage Doucette (KX620751.1) with 97% identity over 64% of the sequence. The PJS22 phage was inserted between the sequence coding for tRNA-Gly(ccc) and a gene encoding DNA protection during starvation protein 2 (PFR_JS22-1_1997), while the PJS23 prophage was inserted between tRNA-Pro(tgg) and a gene encoding Quaternary ammonium compound-resistance protein SugE (PFR_JS23_1469).

Table 11 Phage PJS7 integration study based in PCRs. The fresh reactions were performed at 4 days in three independent experiments from colonies picked randomly from PPA agar plate started directly from glycerol stock. The reactions of propagated cells were performed at 7 days, on cells from colonies generated by re-streaking of the previously tested colony on a new PPA agar plate. All three types of reactions were performed on the cells from the same colony. Presence of all three types of genomes in each colony is consistent with the results of the sequencing study. After 5 generations, the circular phage is no longer observed, while genomes with integrated phage and with no phage at all are present in nearly all tested colonies.

Experiment	Genome without phage	Genome with phage	Circular phage
Fresh			
I	6/6	5/6	5/6
II	7/8	7/8	8/8
III	8/8	8/8	4/8
Propagated			
1st generation	1/8	8/8	7/8
5th generation	8/8	8/8	0/8
10th generation	7/8	8/8	0/8

All of the bacterial genomes were checked for additional prophage sequences with two dedicated programs: Phaster and Prophinder. From the candidate prophages, only the prophage identified from strain JS17 appeared complete. The prophage JS17 was located between the tRNA-Ser(tga) gene and a transposase (PFR_JS17-1_2095). A BLAST search revealed 96% identity over 61% and 64% of the sequence to Propionibacterium phages Doucette and G4, respectively. Similarly, a BLAST search against phages PJS22 and PJS23 showed 97% identity over 62% and 65% of the sequence, respectively.

Based on sequence similarity, the sequenced *P. freudenreichii* phages can be divided into three groups. Group 1 consisted of B22 (KX620750), Doucette (KX620751), E6 (KX620753), G4 (KX620754), Anatole (KX620748), B3 (KX620749), E1 (KX620752), PJS22, PJS23 and the prophage of strain JS17. Group 2 consisted of phages PFR1 (KU984979), PFR2 (KU984980) and PJS7. The one-member Group 3 consisted of the filamentous phage B5 (AF428260), which had a stronger resemblance with the phages that infected Gram-negative bacteria (Chopin *et al.*, 2002) (Figure 22).

Transposases were found in the prophage sequence of strain JS17 and the bacteriophage PJS23. The transposases PFR_JS17-1_2038 and PFR_JS17-1_2067 were both identical to eight (PFR_JS17-1_341, PFR_JS17-1_394, PFR_JS17-1_676, PFR_JS17-1_711, PFR_JS17-1_2205, PFR_JS17-1_2341, PFR_JS17-1_2347 and PFR_JS17-1_2416) and six other transposases (PFR_JS17-1_13, PFR_JS17-1_46, PFR_JS17-1_72, PFR_JS17-1_657, PFR_JS17-1_658 and PFR_JS17-1_1466) found in the same strain, respectively. Additionally, the transposases PFR_JS17-1_657 and PFR_JS17-1_658 were observed as duplicated in part of the genome sequences of strain JS17. These transposases were identical to the transposases found in part of the genome sequences of strains JS12 and JS15 (see Table 8). In phage PJS23,

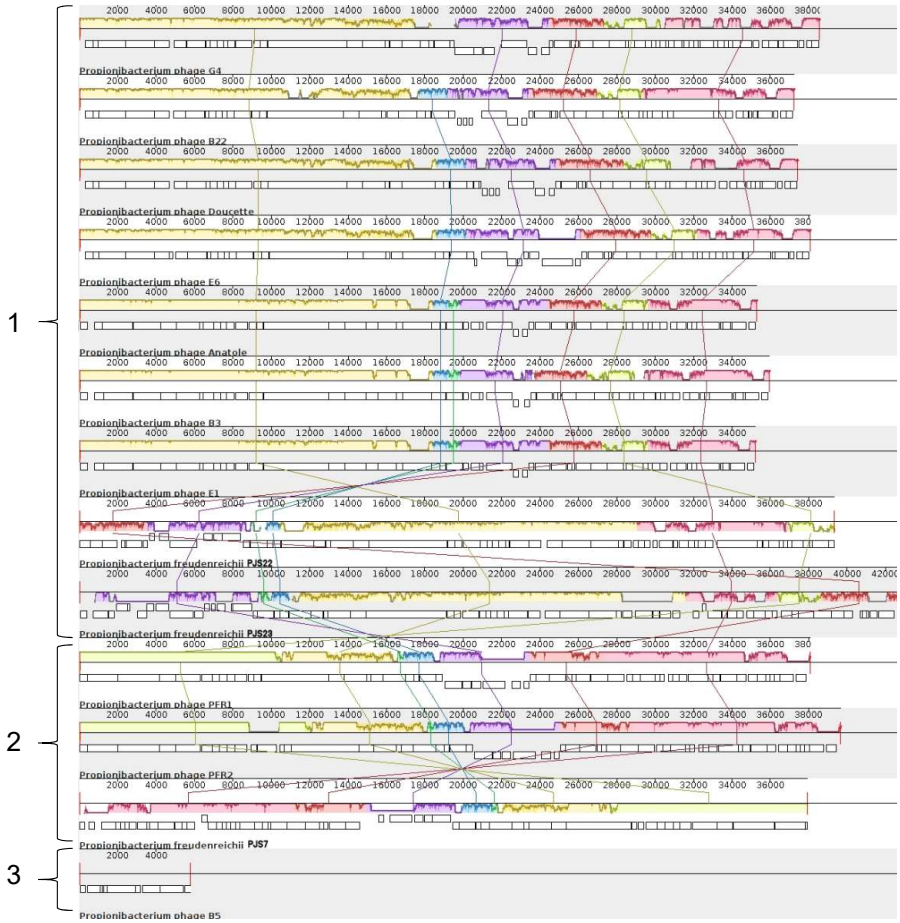

Figure 22 Progressive Mauve genome alignments of all of the sequenced bacteriophages infecting Propionibacterium freudenreichii. Phages form 3 distinct groups. Group: 1 G4, B22, Doucette, E6, Anatole, B3, E1, PJS22, PJS23; Group2: PFR1, PFR1, PJS7; Group 3: filamenous phage B5.

there were four transposase-like elements, PH1_40-PH1_43, co-located on the phage genome (corresponding to PFR_JS23_1432 -PFR_JS23_1435 on the bacterial chromosome). The transposases PFR_JS23_1432 and PFR_JS23_1435 were both found only in the phage region of the genome, while PFR_JS23_1433 (integrase) and PFR_JS23_1434 (transposase) were each found in two additional, co-located copies (PFR_JS23_368 and PFR_JS23_369; PFR_JS23_2185 and PFR_JS23_2184). The transposase pair located on the PJS23 phage genome appeared to be disrupting transposases PFR_JS23_1432 and PFR_JS23_1435. This could mean that the insertion of the new mobile element, intrinsic to the strain, serves as a strategy for better control of the phage and further supports the role of *P. freudenreichii* transposases in the adaptation to a changing environment.

Lysogeny was previously described in *P. freudenreichii* (Herve *et al.,* 2001). However, this is the first study to show *P. freudenreichii*-infecting bacteriophages as complete prophages and as free, circular genomes. The

recent report of the bacteriophage PFR1, which shares 99% sequence identity with the PJS7 phage reported here, demonstrated the phage's ability to enter the lytic cycle and its ability to infect a strain of *Cutibacterium acnes* (Brown *et al.,* 2016). Further studies are needed to establish the dynamics of the relationships between the *P. freudenreichii* strains and the bacteriophages that can infect them.

5.2.4.2 Plasmid-like elements

Owing to the whole genome sequencing from the long reads, two additional circular elements, PFRJS12-3 (LT604882) and PFRJS25-1 (LT618784), were discovered in *P. freudenreichii* strains JS12 and JS25. These elements were of 24.9 kbp and 35.6 kbp in size and encoding 32 and 46 predicted open reading frames, respectively (Figure 23). A BLASTn search revealed that there was no significant similarity between either the previously sequenced plasmids of *P. freudenreichii* (Perez-Chaia *et al.,* 1988; Rehberger and Glatz 1990; Dasen 1998; Kiatpapan *et al.,* 2000; Jore *et al.,* 2001; Stierli 2002), which included p545 (AF291751.1), pRGO1 (NC_002611.1), pLME106 (NC_005705.1) and pLME108 (NC_010065.1), or the *C. acnes* plasmids, which included HL096PA1 (CP003294.1) (Kasimatis *et al.,* 2013) and PA_15_1_R1 (CP012356.1). A BLASTn search of the JS12 plasmid-like element revealed that the gene encoding transposase located at position 8594-9669 had 95% identity to the transposase genes from *P. freudenreichii, A. acidipropionici,* and *Micrococcus luteus* and 93% identity to the transposase gene from *Corynebacterium variabile*. The transposase at position 5799-7130 had 90% sequence identity to *A. acidipropionici* and *M. luteus*, and the transposase at position 12361-12930 had 92% identity to resolvase-encoding genes from *A. acidipropionici*. A BLASTn search of the JS25 plasmid-like element revealed 88% identity to the *Propionibacterium* phage PFR1 over the stretch of its PFR1_23, PFR1_24 and PFR1_25 genes that encoded three hypothetical proteins. Additionally, the 5' end of the sequence showed 98% identity with a 47-nt stretch of the sequence in the non-coding region from *Burkholderia pyrrocinia* strain DSM 10685 plasmid p2327 and the *Burkholderia cenocepacia* J2315 pBCJ2315 plasmid. A BLASTp search against these plasmid sequences showed a negligible sequence similarity.

Further analysis showed that the JS25 sequence was 99% identical over 31% of the sequence of JS12 (Figure 23). The analysis against the Conserved Domains Database (CDD) (Marchler-Bauer *et al.,* 2017) revealed multiple regions of similarity to conjugative plasmids shared between the elements. These regions of similarity included genes coding for the following: a conjugal transfer protein with conserved domains for TrwC relaxase and a TraA exonuclease, which assist in preparing the plasmid DNA for conjugal transfer; TraM, a recognition site for TraD and TraG, which are proteins thought to

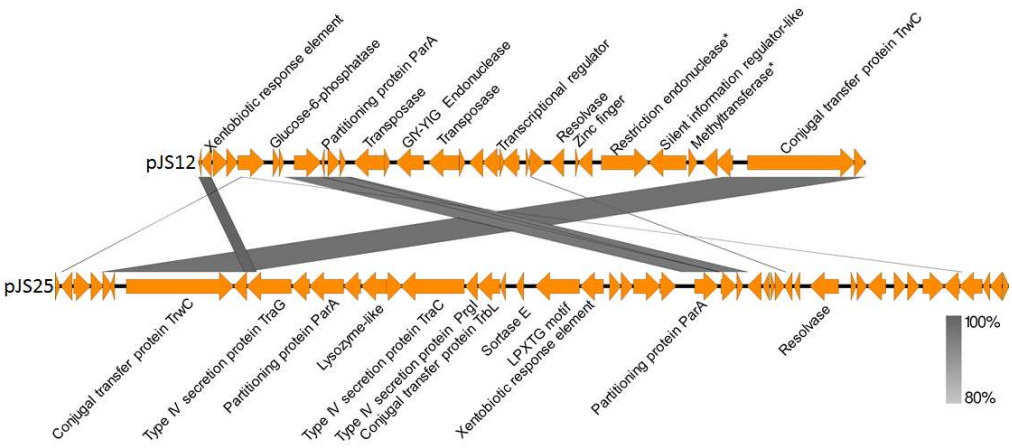

Figure 23 Putative conjugative plasmids identified in study V. The annotations are derived from Conserved Domains database. *The Restriction endonuclease and Methylstransferase marked with asterisk belong to type II restriction-modification system identified in this work.*

mediate interactions between the DNA-processing (Dtr) and the mating pair formation (Mpf) systems; a putative plasmid partition protein, ParA. Based on the presence of these functional genes, the elements could be considered conjugative plasmids. However, it remains to be elucidated whether the circular elements found in strains JS12 and JS25 are plasmids as no characteristic replication origin loci were found. Nevertheless, the presence of transposases with high sequence identity to other known Actinobacteria could mean that these elements are widespread and interchanging among the phylum.

It is noteworthy that the JS25 plasmid-like element appears to be widespread at least among the *P. freudenreichii* strains (Figure 24), because a

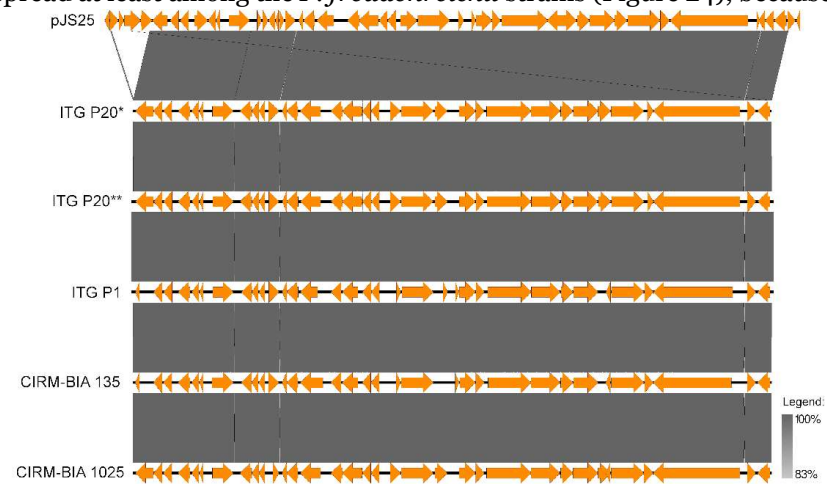

Figure 24 Contigs with 99% identity over 90% of the nucleotide sequence of the JS25 plasmid-like structure in the P. freudenreichii strains ITG P20 (CIRM-BIA 129) (NZ_HG975461 and CCBE010000014**), ITG P1 (CCYV01000031), CIRM-BIA 135 (CCYZ01000006) and CIRM-BIA 1025 (NZ_CCYV01000031).*

comparison with the recently published draft genomes (Falentin *et al.,* 2016; Loux *et al.,* 2015) revealed the presence of sequences with 99% identity over the entire nucleotide sequence in the contigs in *P. freudenreichii* strains ITG P20 (CIRM-BIA 129), ITG P1, CIRM-BIA 135 and CIRM-BIA 1025, constituting 90% of the sequence of the putative plasmid. Because of the limited size of the contigs, it is impossible to determine whether the elements were circular or existed as a part of the chromosome in these strains.

We explored the possibility that the circular elements are types of Integrative and Conjugal Elements (ICEs), which are widely distributed mobile genetic elements. ICEs differ from plasmids by their usual existence within the host's chromosome, but under certain conditions, they can become activated, excise from the chromosome and transfer to a new recipient (Delavat *et al.,* 2016). Although no such elements have been described in Propionibacteria to date, they are widespread among other Actinobacteria as two types. The FtsK/SpoIIIE-type relies on a single FtsK/SpoIIIE-like protein for DNA translocation, while the T4SS-type requires the assembly of a complex type IV translocation system for mobility (Ghinet *et al.,* 2011). Elements that share structural similarities indicative of type T4SS ICEs (Tra and TrwC) are present in putative plasmids pJS12 and pJS25 and are also found in the chromosomes of strains JS7, JS9, JS12, JS18 and JS20. It is therefore possible that the observed plasmid-like elements are mobilised ICEs type T4SS. However, further studies are necessary to determine their true nature.

5.2.4.3 Genomic islands

The genomes were assessed for the presence of genomic islands using IslandViewer 3 (Dhillon *et al.,* 2015), an integrative online tool that performs the analysis with three independent genomic island prediction methods: IslandPick, IslandPath-DIMOB, and SIGI-HMM.

The ability to utilise lactose, a historically important trait in *P. freudenreichii*, has been previously tied to a genomic island in which genes coding for UDP-glucose 4-epimerase (*galE1*), Sodium galactoside symporter (*galP*) and Beta-galactosidase (*lacZ*) are located (Falentin *et al.,* 2010). In the current study, the same island was found in nine other strains: JS, JS2, JS7, JS8, JS10, JS17, JS18, JS22 and JS23, with JS23 possessing two copies of the region (PFR_JS23_160-PFR_JS23_162 and PFR_JS23_2069 - PFR_JS23_2071). Another feature, which is potentially important from the perspective of the dairy industry, is the ability to degrade D-lactate. Eight strains, including JS2, JS7, JS8, JS10, JS15, JS17, JS20 and JS23 had a D-lactate dehydrogenase located on a genomic island, while D-lactate dehydrogenase in strain JS18 was located just downstream of a predicted genomic island. In strain JS4, a genomic island with an alternative pathway for the biosynthesis of rhamnose was found. It consisted of dTDP-4-dehydrorhamnose reductase (*rmlD*), putative dTDP-4-dehydrorhamnose 3,5-epimerase (*rfbC*) and dTDP-glucose 4,6-dehydratase (*rmlB*). In strains JS, JS2, JS4, JS10, JS17, JS18 and JS23, an island with two S-layer protein A genes

was identified. Finally, an island where genes encoded pilus components was found in strain JS18, including Sortase SrtC1 (PFR_J18_2247), Type-2 fimbrial major subunit (PFR_J18_2248) and a Surface-anchored fimbrial subunit (PFR_J18_2249).

The genomes were surveyed for known antibiotic resistance genes, and as previously reported (Thierry *et al.,* 2011), in most cases these genes were not located on a predicted genomic island. However, strain JS8 appeared to be diverging from the group, as the genomic island unique for this strain included three genes that encoded putative antibiotic resistance related proteins: mitomycin radical oxidase (PFR_JS8_1331), tetracycline repressor domain-containing protein (PFR_JS8_1332), and the puromycin resistance protein Pur8 (PFR_JS8_1333). Furthermore, the edges of the island were flanked by the gene coding for hypothetical proteins that have 98 and 99% sequence identity to genes from *Brevibacterium linens* strain SMQ-1335 that encode mobile element proteins.

Additional examples include an island with genes coding for a CASCADE-like CRISPR-Cas system in strains JS2, JS7 and JS9 (see below), an island encoding multiple heat-shock proteins in strains JS9 and JS20, multiple RM systems (see below) and the previously mentioned ICE-like elements in strains JS7, JS9, JS12, JS18 and JS20.

5.2.4.4 CRISPR-Cas systems

The CRISPR together with Cas proteins, commonly known as CRISPR-Cas systems, form the adaptive immunity systems that protect their hosts against invasion by foreign DNA. The function of the adaptive immunity system can be divided into two phenomena: the CRISPR adaptation and CRISPR interference. The CRISPR adaptation results from spacer acquisition upon exposure to invading DNA, while the CRISPR interference involves recognition of the specific spacers on foreign DNA, which in turn allows for the introduction of breaks into those regions of DNA and its resulting destruction (reviewed by Savitskaya *et al.,* 2016). Currently, the CRISPR-Cas systems are subdivided into two classes, five types, and 16 subtypes. The *P. freudenreichii* genomes were analysed with a CRISPR identifying tool, CRISPR-finder. The identified CRISPRs are summarised in Table 12.

After the CRISPR arrays were identified, their vicinities were checked for the presence of the Cas genes. Following the current classification, two types of CRISPR-Cas systems in *P. freudenreichii* were identified, both categorized as class 1, type I CRISPR-Cas systems based on the presence of the typical Cas3 protein (Savitskaya *et al.,* 2016) (Figure 25).

The first system, with the GGATCACCCCCGCGTATGCGGGGAGAAC direct repeat consensus sequence, may be classified as subtype IE based on the sequence homology and gene organisation corresponding to the CASCADE system, which is well-characterised in *E. coli* (Jore *et al.,* 2011; Koonin and Makarova 2013; Wiedenheft *et al.,* 2011). The second system, with the ATTGCCCCTCCTTCTGGAGGGGCCCTTCATTGAGGC direct repeat

consensus sequence bears similarity to the subtype IU (previously IC) (Koonin and Makarova 2013). This classification is strengthened by the presence of the Cas4/Cas1 fusion protein that is found in several variants of the subtype IU (Makarova *et al.,* 2011; Makarova *et al.,* 2015). However, the atypical gene organisation suggests that it is a new variant of the subtype IU.

The CRISPR-Cas system IE was found in strains JS2, JS9 and JS7 and carried 96, 65 and 105 spacers, respectively (Table 12). These systems were located on genomic islands in all strains, suggesting a relatively recent acquisition. However, a lack of sequence identity between the spacers suggests an independent acquisition of immunity (adaptation) in each strain. The CRISPR-Cas system of strain JS7 has a transposase inserted between the *cse1* and *cse2* genes and only a fragment of the *cas2* gene comprising a part of a larger hypothetical protein (Figure 25). In strain JS9, the first 9 CRISPR spacers are separated from the subsequent 96 spacers by an integrase gene. A BLAST search of the spacers indicated immunity to all of the sequenced phages that infect *P. freudenreichii,* except for the filamentous phage B5 for which immunity was found only in strain JS9. Additionally, strain JS2 carried immunity against all three phages found in this study and to the putative plasmid pJS25. Strain JS7 carried immunity to putative plasmid pJS12, phage PJS22 and to the phage by which it was infected, suggesting that the presence of either the transposase or the incomplete *cas2* gene may have resulted in the inactivity of the CRISPR-Cas system in this strain. JS9 carried markers of immunity against all the three phages found in this study and to the pJS25 on the CRISPR_2 array.

The CRISPR-Cas system IU is more widespread in *P. freudenreichii,* and its presence has been previously reported in *P. freudenreichii* strains JS and JS15 (Falentin *et al.,* 2010; Ojala *et al.,* 2017). However, no classification was attempted then. In the current study, the IU system was found in 13 of the strains, including only one strain of cereal origin: JS12. In strains JS4, JS16, JS20, JS21 and JS25 the systems were located on genomic islands. In strains JS11, JS13, JS14, JS18 and JS22 no complete CRISPR-Cas systems were found, although strain JS22 had a short array of CRISPR. No immunity to known phages was found on that array. The number of spacers in IC CRISPR-Cas systems ranged from 25 in strain JS2 to 64 in strain JS17.

The spacers of the strains JS and JS10 and of strains JS4, JS20, JS21 and JS25 were for the most part identical, which is consistent with their phylogenetic relatedness. In other strains, few spacers were identical, suggesting early diversification. Only strain JS2 carried both types of the CRISPR systems, although strain JS9 possessed an additional stretch of 83 spacers separated by a distinct repeat sequence (GGGCTCACCCCCGCATATGCGGGGAGCAC), indicating they belonged to a separate CRISPR-Cas system. Nevertheless, the lack of Cas genes in the vicinity and location of the CRISPR array on the genomic island may mean that the system was acquired in an incomplete form through horizontal gene transfer.

Table 12 Summary of the CRISPR-Cas systems of P. freudenreichii strains

Strain	Predicted CRISPR	Start	End	Size	Direct repeat size	No. of spacers	Type	Conferred immunity to:		Note
								Sequenced phages	Mobile elements found in this study	
JS	Crispr_2	2424099	2427021	2922	36	40	IU	Anatole, B3, B22, E1, E6, G4, PFR1, PFR2	PJS22, PJS23, PJS7, pJS12	-
JS2	Crispr_3	2082550	2088376	5826	28	95	IE	B22, B3, E1, E6, G4, Anatole, Doucette, PFR1, PFR2	PJS22, PJS23, PJS7, pJS25	-
	Crispr_4	2120759	2122610	1851	36	25	IU	PFR1, PFR2	PJS7	-
JS4	Crispr_3	1825170	1827746	2576	36	35	IU	Doucette, E6, G4, PFR1, PFR2	PJS22, PJS7, pJS25	One system, negative strand, spacers separated by a transposase and then by a four-gene integron.
	Crispr_2	1823162	1823697	535	36	7		-	-	
	Crispr_1	1818598	1819215	617	36	8		-	-	
JS7	Crispr_2	2203337	2207335	3998	28	65	IE	B22, B3, E1, E6, G4, Anatole, Doucette, PFR1, PFR2	PJS22, PJS7, pJS12	No *cas2*
JS8	Crispr_3	2192257	2196491	4234	36	58	IU	Anatole, Doucette, B3, B22, E1, E6, G4, PFR1, PFR2, B5	PJS22, PJS23, PJS7, pJS25	-
JS9	Crispr_3	2260079	2260655	576	28	9	IE	-	-	One system, negative strand, spacers separated by an integrase
	Crispr_2	2253046	2258937	5891	28	96		B22, B3, E1, E6, G4, Anatole, Doucette, PFR1, PFR2, B5	PJS22, PJS23, PJS7, pJS25	
	Crispr_4	2288755	2292746	3991	29	65	N/A	Doucette, B3, B22, G4, PFR1, PFR2	PJS23, PJS7	Remnant of a system, no Cas. Spacers separated by an integrase.
	Crispr_5	2293887	2295016	1129	29	18		PFR1, PFR2	PJS7	
JS10	Crispr_2	2029533	2032236	2703	36	37	IU	Anatole, B3, B22, E1, E6, G4, PFR1, PFR2	PJS22, PJS23, PJS7, pJS12	
JS12	Crispr_1	2181462	2185411	3949	36	54	IU	Doucette, B3, B22, E6, G4, PFR1, PFR2, B5	PJS22, PJS23, PJS7	
JS14	PossibleCrispr_1	1595184	1595290	106	36	1	N/A	Anatole, E1		

Strain	Predicted CRISPR	Start	End	Size	Direct repeat size	No. of spacers	Type	Conferred immunity to:		Note
								Sequenced phages	Mobile elements found in this study	
JS15_1	Crispr_1	441841	444266	2425	36	33	IU	Anatole, B3, B22, E1, E6, G4, B5	PJS23, pJS12	
JS16	Crispr_2	417820	419522	1702	36	23	IU	B22, E6, G4, PFR1, PFR2	PJS22, PJS23, PJS7	One system, negative strand, spacers separated by a transposase. Both CRISPR stretches carry immunity to known phages.
	Crispr_1	413342	416347	3005	36	41		Anatole, B3, B22, E1, E6	PJS22	
JS17_1	Crispr_2	2220735	2225402	4667	36	64	IU	Anatole, Doucette, B3, B22, E1, PFR1, PFR2	PJS22, PJS23, pJS12, pJS25	-
JS20_1	Crispr_3	2034463	2037183	2720	36	37	IU	Doucette, B22, E6, G4, PFR1, PFR2	PJS23, PJS7	One system, spacers separated by a transposase and then by an eight-gene integron.
	Crispr_4	2038656	2039048	392	36	5		-	-	
	Crispr_5	2044247	2044936	689	36	9		-	-	
JS21_1	Crispr_5	2189758	2192624	2866	36	39	IU	Doucette, E6, G4, PFR1, PFR2	PJS22, PJS7, pJS25	One system, negative strand, the spacers separated by a transposase and then by a four-gene integron.
	Crispr_4	2187750	2188285	535	36	7		-	-	
	Crispr_3	2183185	2183803	618	36	8		-	-	
JS22	Crispr_2	2150779	2151318	539	36	7	IU	-	-	Remnant of the same IU system, but no Cas. No immunity to known phages found.
JS23	Crispr_2	2124520	2126512	1992	36	27	IU	Anatole, B3, B22, E1, E6, G4, B5	PJS23	-
JS25	Crispr_2	2042081	2045018	2937	36	40	IU	Doucette, E6, PFR1, PFR2	PJS22, PJS7, pJS25	One system, spacers separated by a transposase and then by a four-gene integron.
	Crispr_3	2046491	2046954	463	36	6		-	-	
	Crispr_4	2050901	2051519	618	36	8		-	-	

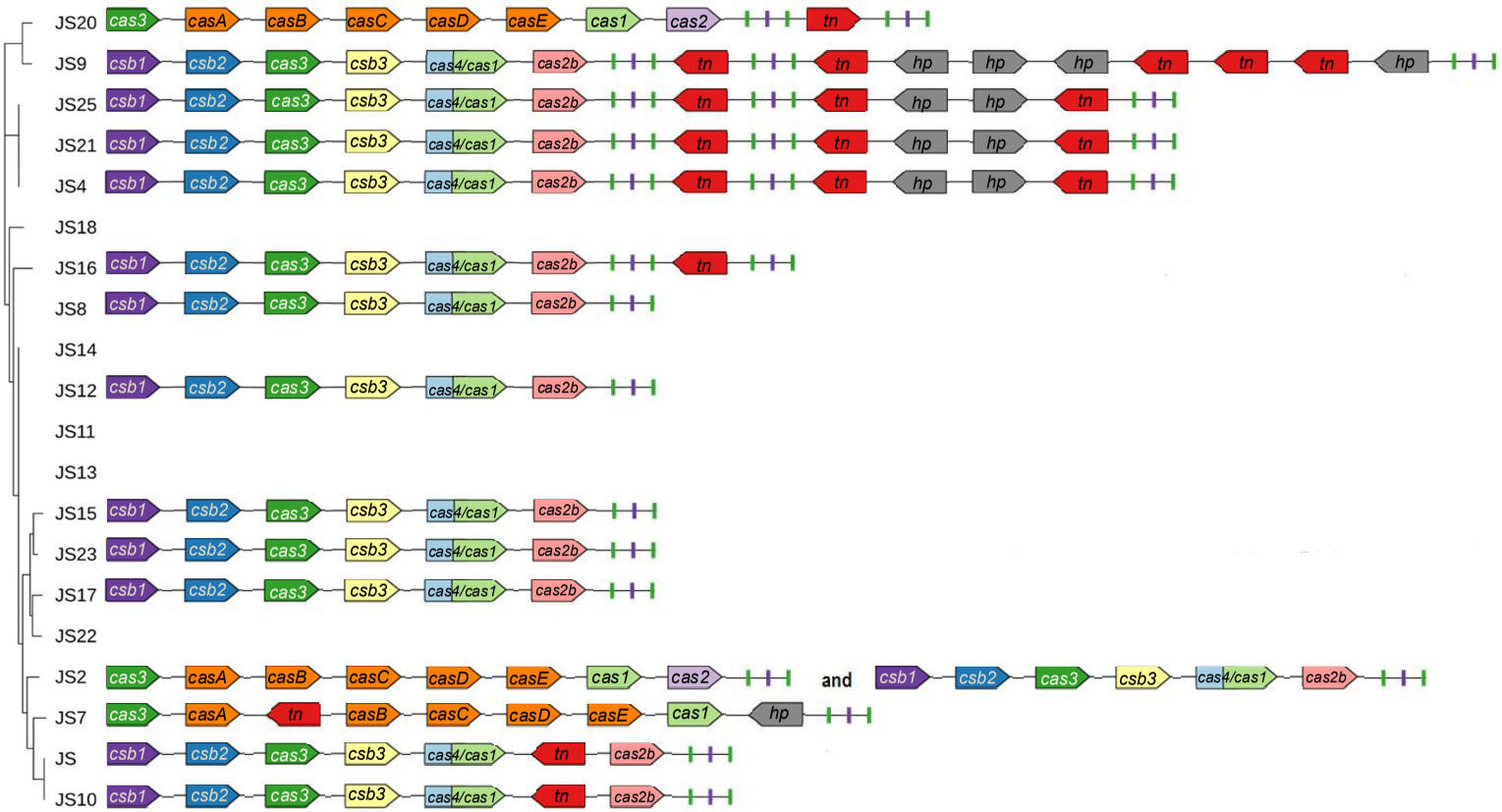

Figure 25 CRISPR-Cas systems in the sequenced strains. Strains JS9, JS2 and JS7 possess CRISPR-Cas system type IE (CASCADE), while all the other strains possess type IU. Only strain JS2 possessess both types of the CRISPR-Cas systems. Color-coding of the genes: green-cas3, orange- CASCADE elements casA, casB, casD, casD, casE; light green- cas1 lilac-cas2, red-transposase/integrase, gray- hypothetical protein, purple- csb1, blue- csb2, yellow-csb3, light blue/green- fusion cas4-cas1, pink- cas2b. Green-purple-green boxes indicate presence of repeats and spacers.

For each of the strains, 2-4 additional "Possible CRISPRs" were identified, most of which mapped within the sequences coding for DNA topoisomerase 1, a hypothetical protein, or fell between coding sequences. None of them showed homology to known Cas genes. Nevertheless, the "Possible CRISPR1" from strain JS14, located within a hypothetical protein coding sequence (PFR_JS14_1397), carried a spacer with 100% identity to a fragment of the gene coding for a tape measure protein in phages Anatole and E1. The significance of this finding remains to be elucidated.

5.2.4.5 Restriction-modification systems

Very little is known about RM systems in *P. freudenreichii*. It has been shown that such systems are present in *P. freudenreichii* based on the observed interference with transformation efficiencies (Jore *et al.* 2001) and the host range that is dependent on the compatibility of the RM systems between the sources and the intended hosts of the plasmids (Kiatpapan *et al.* 2000). The use of the PacBio sequencing platform allows identification of methylation patterns, which in turn permits identification of the recognition motifs of the methylases responsible (Flusberg *et al.,* 2010). The recognition motifs were therefore reported for both strains sequenced with PacBio platform, JS and JS16 (Study IV, Koskinen *et al.,* 2015; Ojala *et al.,* 2017), but the RM systems were not previously characterised.

To gain insight into the possible RM systems present in the 20 strains studied here, the genome sequences were first analysed for the presence of genes that could be identified as components of RM systems. As a result, 216 different RM system genes associated with 127 different systems were identified. The methylases were then assigned to motifs identified by PacBio sequencing. Of the 54 motifs identified, 48 were unambiguously matched to the responsible methylases (Table 13). Among the strains, JS2 and JS7 had three Type I systems, while six strains (JS8, JS9, JS11, JS13, JS14 and JS20) had two such systems and nine strains had a single system. In all of these strains, except JS10, the gene responsible for restriction was intact and the level of methylation was close to complete, suggesting that the systems were active as RM systems.

In addition to the Type I systems, examples of Type II methylases were found in twelve of the strains, including systems with the same specificity in several strains, for instance seven strains contained a methylase recognizing CCWGG motif. One of the Type II systems with the recognition sequence CTCGAG was found on the putative plasmid, pJS12. One Type III RM system, with the unique recognition sequence AGCAGY, was found in five of the strains (JS, JS15, JS17, JS22 and JS23). Overall, the most striking feature of the RM systems distributed among these strains was the variability of systems and genomic locations found from one strain to another. This is in contrast to the more usual situation where at the very least there are one or more common methylases found in all strains of a particular species (see REBASE, Roberts *et al.,* 2015).

Table 13 Methylation motifs and the responsible methylases identified in P. freudenreichii. Recognition sequences in parentheses indicates the methylase responsible cannot be assigned unambiguously. Modified bases are marked in gray; "not annotated" indicates there is no locus_tag as the gene is not annotated in the GenBank file.

Organism	locus	Enzyme	Recognition sequence	Type
P. freudenreichii JS	921	M.PfrJS1II	AGTNNNNNNCTT	I
	566	M.PfrJS1I	AGCAGY	III
P. freudenreichii JS2	PFR_JS2_249	M.PfrJS2I	ATCNNNNNGTCG	I
	PFR_JS2_1297	M.PfrJS2II	AGTNNNNNNCTT	I
	PFR_JS2_2113	M.PfrJS2III	GATNNNNNNTTT	I
	PFR_JS2_2250	M.PfrJS2V	GCCGGC	II
	PFR_JS2_1977	PfrJS2IV	CTTCNAC	
P. freudenreichii JS4	no hits			
P. freudenreichii JS7	PFR_JS7-1_251	M.PfrJS7I	ATCNNNNNGTCG	I
	PFR_JS7-1_659	M.PfrJS7II	CAANNNNNNCTG	I
	PFR_JS7-1_2249	M.PfrJS7III	CGANNNNNNNTTT	I
	PFR_JS7-1_2117	M.PfrJS7IV	CCWGG	II
	PFR_JS7-1_1519	M.PfrJS7V	CTGAAG	II
P. freudenreichii PJS7 phage	PFR_JS7-PH_14	M1.PfrJS7aORF14P	(not annotated)	II
	PFR_JS7-PH_22	M2.PfrJS7aORF14P	(not annotated)	II
P. freudenreichii JS8	PFR_JS8_242	M.PfrJS8I	ATCNNNNNGTCG	I
	PFR_JS8_2157	M.PfrJS8III	CAGNNNNNNRTCG	I
	PFR_JS8_2013	M.PfrJS8II	CCWGG	II
P. freudenreichii JS9	PFR_JS9-1_274	M.PfrJS9I	GCANNNNNNGTGG	I
	PFR_JS9-1_1487	M.PfrJS9II	CAGNNNNNNCTG	I
	PFR_JS9-1_705	M.PfrJS9III	CTGGAG	II
	PFR_JS9-1_1590	M.PfrJS9IV	ARCCCC	
P. freudenreichii JS10	PFR_JS10_1239	M.PfrJS10I	AGTNNNNNNCTT	I
P. freudenreichii JS11	PFR_JS11_224	M.PfrJS11I	ACCNNNNNCTTC	I
	PFR_JS11_2032	M.PfrJS11II	GGANNNNNNNCTC	I
	PFR_JS11_1917	M.PfrJS11III	CCWGG	II
P. freudenreichii JS12	PFR_JS12-1_2109	M.PfrJS12I	GCANNNNNNRTGA	I
	PFR_JS12-1_1387	M.PfrJS12III	GGGCCC (GGCGGAG) (TANAAG)	II
P. freudenreichii pJS12	PFR_JS12-3_27	M.PfrJS12II	CTCGAG	II
P. freudenreichii JS13	PFR_JS13-1_224	M.PfrJS13I	ACCNNNNNCTTC	I
	PFR_JS13-1_2034	M.PfrJS13II	GGANNNNNNNCTC	I
	PFR_JS13-1_1919	M.PfrJS13III	CCWGG	II

Organism	locus	Enzyme	Recognition sequence	Type
P. freudenreichii JS14	PFR_JS14_217	M.PfrJS14I	ATCNNNNNGTCG	I
	PFR_JS14_1997	M.PfrJS14III	CAGNNNNNNCTG	I
	PFR_JS14_1883	M.PfrJS14II	CCWGG	II
P. freudenreichii JS15	PFR_JS15-1_281	M.PfrJS15II	CCWGG	II
	PFR_JS15-1_2040	M.PfrJS15I	AGCAGY	III
	PFR_JS15-1_303	PfrJS15III	CTTCNAC	
P. freudenreichii JS16	RM25_0134	M.Pfr20271I	GGANNNNNNNCTT	I
			(TCGWCGA)	II
P. freudenreichii JS17	PFR_JS17-1_893	M.PfrJS17I*	GAANNNNNNNCTT	I
	PFR_JS17-1_273	M.PfrJS17II	AGCAGY	III
P. freudenreichii JS18	J18_2127	M.PfrJS18I	CAGNNNNNNCTG	I
P. freudenreichii JS20	PFR_JS20-1_520	M.PfrJS20I	GACGTC	
	PFR_JS20-1_402	M.PfrJS20II	ACGNNNNNNGTGG	I
	PFR_JS20-1_2119	M.PfrJS20III	GCANNNNNNGTGG	I
P. freudenreichii JS21	PFR_JS21-1_2269	M.PfrJS21I	ACANNNNNNRTAY	I
	PFR_JS21-1_2026	M.PfrJS21II	CCWGG	II
P. freudenreichii JS22	PFR_JS22-1_2165	M.PfrJS22I**	GAANNNNNNNCTT	I
	PFR_JS22-1_1883	M.PfrJS22II	GRGCYC	II
	PFR_JS22-1_271	M.PfrJS22III	AGCAGY	III
			(CCGCTC)	
P. freudenreichii JS22 phage	PFR_JS22-PH_50	M.PfrJS22PORF50P	(m5C) (not annotated)	II
P. freudenreichii JS23	PFR_JS23_2000	M.PfrJS23I	AGCAGY	III
	PFR_JS23_1965	PfrJS23II	CTTCNAC	
P. freudenreichii JS25	PFR_JS25-1_2023	M.PfrJS25I	ACANNNNNNRTAY	I

*It is possible that that M.PfrJS17ORF2252P is the active methylase in this organism. M.PfrJS17ORF2252P shares 99% aa sequence identity with M.PfrJS22I, but its corresponding S protein is truncated.
**It is possible that M.PfrJS22ORF850P is the active methylase in this organism. M.PfrJS22ORF850P is 100% identical to M.PfrJS17I.

5.2.4.6 Pilus and mucus binding

Pili are surface adhesive components and well-documented virulence factors for many opportunistic pathogens. However, their role in probiotics and/or commensal bacteria as niche-adaptation factors has recently been recognised. While the pili of probiotic Bifidobacteria have been reported as crucial for establishing host-microbe and microbe-microbe interactions

(Turroni *et al.,* 2013; O'Connell Motherway *et al.,* 2011), no pili have been reported in Propionibacteria to date.

A search for pilus clusters using the LOCP tool (Plyusnin *et al.*, 2009) identified putative pilus operons in the genomes of JS18, JS20 and JS14, consisting of three, four and five open reading frames (ORFs), respectively (Figure 26A). The putative pilus operons in JS14 and JS20 were similarly organised, and the predicted pilus proteins share 99-94% amino acid identity. The ORF prediction and location of functional domains in predicted proteins suggested that in the JS14 genome, the genes coding for the putative surface-anchored fimbrial subunit and the Type-2 fimbrial major subunit were split, possibly due to frameshift-causing mutations. A BLASTp search revealed the presence of highly conserved gene clusters in all of the *P. freudenreichii* genomes studied here, with a similar structural organisation to the pilus operons of JS14 and JS20. The exception was strain JS9, where only partial genes coding for surface-anchored fimbrial subunit (PFR_JS9-1_404) and sortase (PFR_JS9-1_414) were found with several genes inserted between them. In contrast, homology searches revealed that the PFR_J18_2249-PFR_J18_2248-PFR_J18_2247 operon that was in a genomic island was unique for the JS18 strain because no counterpart of the intact operon was identified in the other genomes. Per the BLAST searches, the genome of JS9 carried genes encoding the putative surface-anchored fimbrial subunit (PFR_JS9-1_546) and the Type-2 fimbrial major subunit (PFR_JS9-1_547) with 100% identity to the gene products of the JS18 genome, but the third gene encoding the putative sortase was not present. The predicted pilus proteins encoded by the JS18 operon shared 32-54% identity with their counterparts encoded by the JS14/JS20 operon, and a BLASTp search against the NCBI non-redundant protein database revealed the highest sequence identity (39-55%) to proteins from *Haematomicrobium sanguinis*.

Because *in silico* searches suggested that the JS18 and JS20 genomes carried intact pilus operons, these strains were chosen for electron microscopic analyses together with a *P. freudenreichii*-type strain (JS16) as the control. The transmission electron microscopic images of negatively stained cells showed that the surfaces of JS18 cells contained pilus-like appendages, which were not observed in the JS20 and JS16 cells (Figure 26B).

Efficient and specific mucus binding is a feature that has not been described for *Propionibacterium* strains. Because mucus-binding is one well-known function that is mediated by pili in other bacteria (Kankainen *et al.,* 2009) and because electro microscopy revealed pilus-like structures on the JS18 cell surface, specific mucus binding to JS18 was tested and compared to the non-specific binding of the control protein, BSA, to JS18. The test revealed that the JS18 strain adhered more efficiently to the mucus than to BSA, while no specific binding to mucus was observed in the other strains tested (Figure 26C). The presence of pili and their positive effect on the binding capacity of the strain could make JS18 a unique probiotic candidate, and its probiotic potential should be explored further.

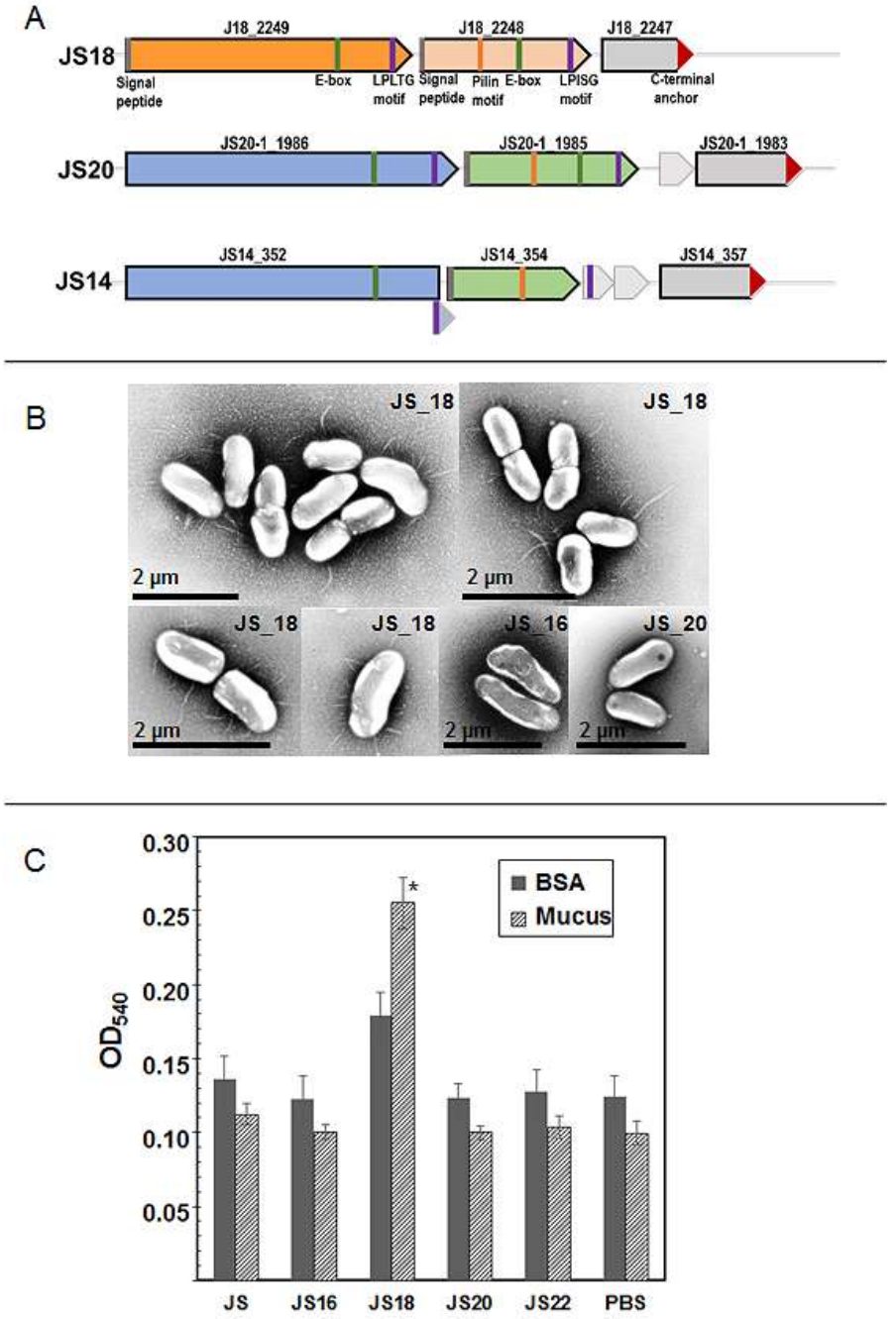

Figure 26 Pili in P. freudenreichii. Panel A: Pili clusters identified by LOCP; Panel B: Transmission electron microscopy images of strains JS18 and JS20 carrying the two types of pili operons identified by LOCP, as well as JS16 strain used as control; Panel C: The results of mucus binding assay. The binding ability of strain JS18 to mucus was significantly higher than non-specific binding to BSA and also significantly higher than the control value obtained with PBS and is marked with an asterisk.

5.2.4.7 Vitamin B_{12} biosynthetic pathway

The vitamin B_{12} biosynthetic pathway in *P. freudenreichii* has been previously resolved (Falentin *et al.*, 2010; Murooka *et al.,* 2005; Murakami *et al.,* 1993; Roessner *et al.,* 2002; Sattler *et al.,* 1995). All of the strains included in this study demonstrated an ability to produce active vitamin B_{12} (Study II), and it was confirmed by BLASTp searches that all of the strains possess the previously identified B_{12} biosynthetic genes, with a similar local organisation into clusters, and with high level of sequence conservation. Strain JS4 was the exception because the *hemL* and *cbiD* genes were shorter, and the gene *cbiX* appeared to be missing. However, it was determined that in this strain, a one-nucleotide-shorter spacer region resulted in a frameshift, which in turn led to the formation of a fusion gene of *cbiX* with the preceding gene *cbiH*. The visual assessment of the raw sequencing data confirmed a deletion of a guanine base in the region, which caused the frameshift. Additionally, the predicted sizes of *cbiX* and *cbiH* had 18- and 15-nucleotide variations between strains, respectively, pointing to the variable character of the spacer region.

The B_{12} biosynthetic pathway is regulated at the translational level by the cobalamin riboswitches (Nahvi *et al.,* 2004). In *P. freudenreichii,* three of these riboswitches have been found upstream of genes *cbiL*, *cbiB* and *mutA* (Vitreschak *et al.,* 2003). The B_{12}-riboswitches in *P. freudenreichii* are not well characterised, and the actual span of the element is unknown. However, all the elements are expected to possess the conserved B_{12} binding region, termed the B_{12}-box, which is characterised by a consensus sequence, rAGYCMGgAgaCCkGCcd (Vitreschak *et al.,* 2003). Based on previous reports (Nahvi *et al.*, 2004; Vitreschak *et al.,* 2003), the predicted sequences for the three putative riboswitches were retrieved and compared among the strains. All of the tested *P. freudenreichii* strains possessed the expected three riboswitches, which were highly conserved between the strains, with the B_{12}-box consensus for the species, SAGYCMSAMRMBCYGCCD. It may be significant that the riboswitches of the *cbiL* and *mutA* genes are located very close to each other and in opposite orientations, but the implications of this positioning remain unclear.

6 CONCLUSIONS

The results presented in this study confirm that *P. freudenreichii* is capable of synthesising endogenous DMBI from reduced FMN, as long as oxygen is available during fermentation. Under strictly anaerobic conditions, no production of active vitamin occurs, which was attributed to the inability to form DMBI by the oxygen-dependent BluB domain of the BluB/CobT2 fusion enzyme.

The yields of vitamin B_{12} produced by the strains that were grown in supplemented WBM medium were strain-dependent. All of the *P. freudenreichii* strains produced active vitamin B_{12} under all conditions tested, while *A. acidipropionici* produced small amounts of vitamin only when DMBI was supplied. The overall high yields in the *P. freudenreichii* strains were attributed to the addition of cobalt and the consumption of high amounts of riboflavin naturally present in the medium. Supplementation with DMBI or nicotinamide had an additional positive effect on vitamin B_{12} yields but only in some of the strains. These results suggest that *P. freudenreichii* does not require DMBI supplementation under the conditions tested.

Finally, it was showed that when grown in two media with different effects on cell metabolism, *P. freudenreichii* 2067 produced nutritionally relevant amounts of active vitamin B_{12} without the need for cobalt or DMBI supplementation. *P. freudenreichii* is a food-grade bacterium, and the results confirm that it can be used for *in situ* fortification of foods in vitamin B_{12}, where addition of neither DMBI nor cobalt is allowed.

Genome sequencing did not reveal a clear genetic background for the differences in the levels of vitamin B_{12} production between the tested strains, but the availability of the genomic data will facilitate future studies. However, the whole genome sequencing from the long reads enabled the discovery of previously unknown characteristics of the species. Among them, the potential transposase-mediated adaptation strategies reflected by the differences in genome organisation between closely related strains was described; two putative-plasmid structures were identified, and for the first time, the complete prophages that integrated into the bacterial chromosomes and were also observed in their free, circular forms were reported. Moreover, the systems that protected *P. freudenreichii* from invading DNA, namely, the CRISPR-Cas and the RM systems were analysed using bioinformatics tools. Both types of the defense systems have been observed in *P. freudenreichii* before, but their classification, diversity and distribution was described for the first time in current work.

Finally, an unprecedented complete pilus structure formed by the *P. freudenreichii* JS18 strain was identified and then associated with the specific mucus binding. Taken together, the results reported here improve the genomic characterisation of *P. freudenreichii* and provide a foundation for further, more in-depth studies of the species.

7 REFERENCES

1. Aburjaile, F.F., Rohmer, M., Parrinello, H., Maillard, M.B., Beaucher, E., Henry, G., Nicolas, A., Madec, M.N., Thierry, A., Parayre, S. and Deutsch, S.M., 2016. Adaptation of *Propionibacterium freudenreichii* to long-term survival under gradual nutritional shortage. *BMC genomics*, *17*(1), p.1007.
2. Allen, R.H. and Stabler, S.P., 2008. Identification and quantitation of cobalamin and cobalamin analogues in human feces. *The American journal of clinical nutrition*, *87*(5), pp.1324-1335.
3. Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research 25*, 3389-3402.
4. Antony, A.C., 2003. Vegetarianism and vitamin B-12 (cobalamin) deficiency. *The American journal of clinical nutrition*, *78*(1), pp.3-6.
5. Arndt, D., Grant, J.R., Marcu, A., Sajed, T., Pon, A., Liang, Y. and Wishart, D.S., 2016. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic acids research*, *44*(W1), pp.W16-W21.
6. Avissar, Y.J., Ormerod, J.G. and Beale, S.I., 1989. Distribution of δ-aminolevulinic acid biosynthetic pathways among phototrophic bacterial groups. *Archives of microbiology*, *151*(6), pp.513-519.
7. Ball, G.F.M. 2006. Chapter 18: Microbiological Methods for the Determination of the B-Group Vitamins in Vitamins in foods: Analysis, bioavailability and stability. *Boca Raton: CRC Press*.
8. Banerjee, R. and Ragsdale, S.W., 2003. The many faces of Vitamin B12: catalysis by cobalamin-dependent enzymes 1. *Annual review of biochemistry*, *72*(1), pp.209-247.
9. Battersby, A.R. and Leeper, F.J., 1999. Biosynthesis of B12 in the aerobic organism *Pseudomonas denitrificans*. *Chemistry and Biochemistry of B12*, pp.507-535.
10. Berry, E.C. and Bullerman, L.B., 1966. Use of Cheese Whey for Vitamin B12 Production II. Cobalt, Precursor, and Aeration Levels. *Applied microbiology*, *14*(3), pp.356-357.
11. Bhatia, S. and Densmore, D., 2013. Pigeon: a design visualizer for synthetic biology. *ACS synthetic biology*, *2*(6), pp.348-350.
12. Bhushan, B., Tomar, S.K. and Chauhan, A., 2017. Techno-functional differentiation of two vitamin B12 producing *Lactobacillus plantarum* strains: an elucidation for diverse future use. *Applied Microbiology and Biotechnology*, *101*(2), pp.697-709.
13. Black, M.M., 2008. Effects of vitamin B12 and folate deficiency on brain development in children. *Food and nutrition bulletin*, *29*(2_suppl1), pp.S126-S131.
14. Blanche, F., Cameron, B., Crouzet, J., Debussche, L., Thibaut, D., Vuilhorgne, M., Leeper, F.J. and Battersby, A.R., 1995. Vitamin B12: how the problem of its biosynthesis was solved. *Angewandte Chemie International Edition in English*, *34*(4), pp.383-411.
15. Brown, T.L., Tucci, J., Dyson, Z.A., Lock, P., Adda, C.G. and Petrovski, S., 2017. Dynamic interactions between prophages induce lysis in *Propionibacterium acnes*. *Research in microbiology*, *168*(2), pp.103-112.
16. Bullerman, L.B. and Berry, E.C., 1966a. Use of Cheese Whey for Vitamin B12 Production I. Whey Solids and Yeast Extract Levels. *Applied microbiology*, *14*(3), pp.353-355.
17. Bullerman, L.B. and Berry, E.C., 1966b. Use of Cheese Whey for Vitamin B12 Production III. Growth Studies and Dry-Weight Activity. *Applied microbiology*, *14*(3), pp.358-360.
18. Burgess, C.M., Smid, E.J. and van Sinderen, D., 2009. Bacterial vitamin B2, B11 and B12 overproduction: an overview. *International journal of food microbiology, 133*(1), pp.1-7.
19. Capozzi, V., Russo, P., Dueñas, M.T., López, P. and Spano, G., 2012. Lactic acid bacteria producing B-group vitamins: a great potential for functional cereals products. *Applied microbiology and biotechnology, 96*(6), pp.1383-1394.
20. Carmel, R., 2008. How I treat cobalamin (vitamin B12) deficiency. *Blood*, *112*(6), pp.2214-2221.
21. Chamlagain, B., Edelmann, M., Kariluoto, S., Ollilainen, V. and Piironen, V., 2015. Ultra-high performance liquid chromatographic and mass spectrometric analysis of active vitamin B12 in cells of *Propionibacterium* and fermented cereal matrices. *Food chemistry*, *166*, pp.630-638.
22. Chin, C.S., Alexander, D.H., Marks, P., Klammer, A.A., Drake, J., Heiner, C., Clum, A., Copeland, A., Huddleston, J., Eichler, E.E. and Turner, S.W., 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nature methods*, *10*(6), pp.563-569.
23. Chopin, M.C., Rouault, A., Ehrlich, S.D. and Gautier, M., 2002. Filamentous phage active on the gram-positive bacterium *Propionibacterium freudenreichii*. *Journal of bacteriology*, *184*(7), pp.2030-2033.
24. Collins, H.F., Biedendieck, R., Leech, H.K., Gray, M., Escalante-Semerena, J.C., McLean, K.J., Munro, A.W., Rigby, S.E., Warren, M.J. and Lawrence, A.D., 2013. *Bacillus megaterium* has both a functional

BluB protein required for DMB synthesis and a related flavoprotein that forms a stable radical species. *PloS one*, *8*(2), p.e55708.
25. Commission Regulation (EU) No 1129/2011 of 11 November 2011 amending Annex II to Regulation (EC) No 1333/2008 of the European Parliament and of the Council by establishing a Union list of food additives Text with EEA relevance, OJ L 295, 12.11.2011, p. 1–177
26. Crofts, T.S., Seth, E.C., Hazra, A.B. and Taga, M.E., 2013. Cobamide structure depends on both lower ligand availability and CobT substrate specificity. *Chemistry & biology*, *20*(10), pp.1265-1274.
27. Crow, V.L., 1986. Utilization of lactate isomers by *Propionibacterium freudenreichii* subsp. *shermanii*: regulatory role for intracellular pyruvate. *Applied and environmental microbiology*, *52*(2), pp.352-358.
28. Dalmasso, M., Nicolas, P., Falentin, H., Valence, F., Tanskanen, J., Jatila, H., Salusjärvi, T. and Thierry, A., 2011. Multilocus sequence typing of *Propionibacterium freudenreichii*. *International journal of food microbiology, 145*(1), pp.113-120.
29. Daly, D.F., McSweeney, P.L. and Sheehan, J.J., 2010. Split defect and secondary fermentation in Swiss-type cheeses–a review. Dairy science & technology, 90(1), pp.3-26.
30. Darling, A.E., Mau, B. and Perna, N.T., 2010. ProgressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PloS one*, *5*(6), p.e11147.
31. Dasen, G.H., 1998. *Molecular identification and applied genetics of propionibacteria* (Doctoral dissertation). ETH Zürich, Nr. 12981, 1998.
32. De Angelis, M., Bottacini, F., Fosso, B., Kelleher, P., Calasso, M., Di Cagno, R., Ventura, M., Picardi, E., van Sinderen, D. and Gobbetti, M., 2014. *Lactobacillus rossiae*, a vitamin B 12 producer, represents a metabolically versatile species within the genus *Lactobacillus*. *PloS one*, *9*(9), p.e107232.
33. De Freitas, R., Madec, M.N., Chuat, V., Maillard, M.B., Mukdsi, M.C.A., Falentin, H., de Carvalho, A.F., Valence, F. and Thierry, A., 2015. New insights about phenotypic heterogeneity within *Propionibacterium freudenreichii* argue against its division into subspecies. *Dairy science & technology*, *95*(4), pp.465-477.
34. Degnan, P.H., Barry, N.A., Mok, K.C., Taga, M.E. and Goodman, A.L., 2014a. Human gut microbes use multiple transporters to distinguish vitamin B 12 analogs and compete in the gut. *Cell host & microbe*, *15*(1), pp.47-57.
35. Degnan, P.H., Taga, M.E. and Goodman, A.L., 2014b. Vitamin B 12 as a modulator of gut microbial ecology. *Cell metabolism*, *20*(5), pp.769-778.
36. Delavat, F., Mitri, S., Pelet, S. and van der Meer, J.R., 2016. Highly variable individual donor cell fates characterize robust horizontal gene transfer of an integrative and conjugative element. *Proceedings of the National Academy of Sciences*, p.201604479.
37. Dherbecourt, J., Bourlieu, C., Maillard, M.B., Aubert-Frogerais, L., Richoux, R. and Thierry, A., 2010. Time course and specificity of lipolysis in Swiss cheese. Journal of agricultural and food chemistry, 58(22), pp.11732-11739.
38. Dhillon, B.K., Laird, M.R., Shay, J.A., Winsor, G.L., Lo, R., Nizam, F., Pereira, S.K., Waglechner, N., McArthur, A.G., Langille, M.G. and Brinkman, F.S., 2015. IslandViewer 3: more flexible, interactive genomic island discovery, visualization and analysis. *Nucleic acids research*, 43, p.104-108.
39. Edelmann, M., Chamlagain, B., Santin, M., Kariluoto, S. and Piironen, V., 2016. Stability of added and in situ-produced vitamin B12 in breadmaking. *Food chemistry*, *204*, pp.21-28.
40. EFSA BIOHAZ Panel (EFSA Panel on Biological Hazards), 2012. Scientific opinion on the maintenance of the list of QPS biological agents intentionally added to food and feed (2012 update), EFSA Journal, vol. 10, no. 12, p. 3020
41. EFSA NDA Panel (EFSA Panel on Dietetic Products, Nutrition and Allergies) (2015). Scientific opinion on dietary reference values for cobalamin (vitamin B12). *EFSA J. 13*:64.
42. Elliott, T. and Roth, J.R., 1989. Heme-deficient mutants of *Salmonella typhimurium*: two genes required for ALA synthesis. *Molecular and General Genetics MGG*, *216*(2), pp.303-314.
43. Eschenmoser, A. and Wintner, C.E., 1977. Natural product synthesis and vitamin B12. *Science*, *196*(4297), pp.1410-1420.
44. Eshel, G., Shepon, A., Noor, E. and Milo, R., 2016. Environmentally Optimal, Nutritionally Aware Beef Replacement Plant-Based Diets. *Environmental Science & Technology*, *50*(15), pp.8164-8168.
45. Falentin, H., Deutsch, S.M., Jan, G., Loux, V., Thierry, A., Parayre, S., Maillard, M.B., Dherbécourt, J., Cousin, F.J., Jardin, J. and Siguier, P., 2010. The complete genome of *Propionibacterium freudenreichii* CIRM-BIA1 T, a hardy Actinobacterium with food and probiotic applications. *PloS one*, *5*(7), p.e11748.
46. Falentin, H., Deutsch, S.M., Loux, V., Hammani, A., Buratti, J., Parayre, S., Chuat, V., Barbe, V., Aury, J.M., Jan, G. and Le Loir, Y., 2016. Permanent draft genome sequence of the probiotic strain

Propionibacterium freudenreichii CIRM-BIA 129 (ITG P20). *Standards in genomic sciences*, *11*(1), p.6.
47. Field, D., Tiwari, B., Booth, T., Houten, S., Swan, D., Bertrand, N. and Thurston, M., 2006. Open software for biologists: from famine to feast. *Nature biotechnology*, *24*(7), p.801.
48. Flusberg,B.A., Webster,D.R., Lee,J.H., Travers,K.J., Olivares,E.C., Clark,T.A., Korlach,J. and Turner,S.W. (2010) Direct detection of DNA methylation during single-molecule, real-time sequencing. Nat. Methods, 7, 461–465.
49. Friedmann, H.C., 1965. Partial purification and properties of a single displacement trans-N-glycosidase. *Journal of Biological Chemistry*, *240*(1), pp.413-418.
50. Ghinet, M.G., Bordeleau, E., Beaudin, J., Brzezinski, R., Roy, S. and Burrus, V., 2011. Uncovering the prevalence and diversity of integrating conjugative elements in actinobacteria. *PLoS One*, *6*(11), p.e27846.
51. Gray, M.J. and Escalante-Semerena, J.C., 2007. Single-enzyme conversion of FMNH2 to 5, 6-dimethylbenzimidazole, the lower ligand of B12. *Proceedings of the National Academy of Sciences*, *104*(8), pp.2921-2926.
52. Green, R. and Miller, J.W., 2013 in Vitamin B12. *Handbook of Vitamins*, Fifth edition, Zempleni, J., Suttie, J.W., Gregory III, J.F., Stoverp, P.J. (ed.) CRC Press, p.447.
53. Grissa, I., Vergnaud, G. and Pourcel, C., 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic acids research*, *35*(suppl 2), pp.W52-W57.
54. Gu, Q., Zhang, C., Song, D., Li, P. and Zhu, X., 2015. Enhancing vitamin B 12 content in soy-yogurt by *Lactobacillus reuteri*. *International journal of food microbiology*, *206*, pp.56-59.
55. Hadfield, J., Croucher, N.J., Goater, R.J., Abudahab, K., Aanensen, D.M. and Harris, S.R., 2017. Phandango: an interactive viewer for bacterial population genomics. *bioRxiv*, p.119545.
56. Hargrove, R.E. and Leviton, A., 1955. *Process for the manufacture of vitamin b12*. U.S. Patent 2,715,602.
57. Hashimoto, Y., Yamashita, M. and Murooka, Y., 1997. The Propionibacterium freudenreichii hemYHBXRL gene cluster, which encodes enzymes and a regulator involved in the biosynthetic pathway from glutamate to protoheme. *Applied microbiology and biotechnology*, *47*(4), pp.385-392.
58. Hashimoto, Y., Yamashita, M., Ono, H. and Murooka, Y., 1996. Characterization of the hemB gene encoding δ-aminolevulinic acid dehydratase from *Propionibacterium freudenreichii*. *Journal of fermentation and bioengineering*, *82*(2), pp.93-100.
59. Hazra, A.B., Han, A.W., Mehta, A.P., Mok, K.C., Osadchiy, V., Begley, T.P. and Taga, M.E., 2015. Anaerobic biosynthesis of the lower ligand of vitamin B12. *Proceedings of the National Academy of Sciences*, *112*(34), pp.10792-10797.
60. Hazra, A.B., Tran, J.L., Crofts, T.S. and Taga, M.E., 2013. Analysis of substrate specificity in CobT homologs reveals widespread preference for DMB, the lower axial ligand of vitamin B 12. *Chemistry & biology*, *20*(10), pp.1275-1285.
61. Hervé, C., Coste, A., Rouault, A., Fraslin, J.M. and Gautier, M., 2001. First Evidence of Lysogeny in *Propionibacterium freudenreichii* subsp. *shermanii*. *Applied and environmental microbiology*, *67*(1), pp.231-238.
62. Hörig, J.A. and Renz, P., 1980. Biosynthesis of Vitamin B12. *European Journal of Biochemistry*, *105*(3), pp.587-592.
63. Hugenschmidt, S., Schwenninger, S.M. and Lacroix, C., 2011. Concurrent high production of natural folate and vitamin B12 using a co-culture process with *Lactobacillus plantarum* SM39 and *Propionibacterium freudenreichii DF13*. *Process Biochemistry*, *46*(5), pp.1063-1070.
64. Hugenschmidt, S., Schwenninger, S.M., Gnehm, N. and Lacroix, C., 2010. Screening of a natural biodiversity of lactic and propionic acid bacteria for folate and vitamin B12 production in supplemented whey permeate. *International dairy journal*, *20*(12), pp.852-857.
65. Huptas, C., Scherer, S. and Wenning, M., 2016. Optimized Illumina PCR-free library preparation for bacterial whole genome sequencing and analysis of factors influencing de novo assembly. *BMC research notes*, *9*(1), p.269.
66. Iida, K. and Kajiwara, M., 2000. Evaluation of biosynthetic pathways to δ-aminolevulinic acid in *Propionibacterium shermanii* based on biosynthesis of vitamin B12 from D-[1-13C] glucose. *Biochemistry*, *39*(13), pp.3666-3670.
67. Iida, K., Ohtaka, K. and Kajiwara, M., 2007. Mechanism of the ring contraction process in vitamin B12 biosynthesis by the anaerobe *Propionibacterium shermanii* under aerobic conditions. *FEBS Journal*, *274*(13), pp.3475-3481.

68. Jore, J.P.M., Van Luijk, N., Luiten, R.G.M., Van der Werf, M.J. and Pouwels, P.H., 2001. Efficient Transformation System for *Propionibacterium freudenreichii* Based on a Novel Vector. *Applied and environmental microbiology*, *67*(2), pp.499-503.
69. Jore, M.M., Lundgren, M., Van Duijn, E., Bultema, J.B., Westra, E.R., Waghmare, S.P., Wiedenheft, B., Pul, Ü., Wurm, R., Wagner, R. and Beijer, M.R., 2011. Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nature structural & molecular biology*, *18*(5), pp.529-536.
70. Kankainen, M., Paulin, L., Tynkkynen, S., von Ossowski, I., Reunanen, J., Partanen, P., Satokari, R., Vesterlund, S., Hendrickx, A.P., Lebeer, S. and De Keersmaecker, S.C., 2009. Comparative genomic analysis of *Lactobacillus rhamnosus* GG reveals pili containing a human-mucus binding protein. *Proceedings of the National Academy of Sciences*, *106*(40), pp.17193-17198.
71. Kasimatis, G., Fitz-Gibbon, S., Tomida, S., Wong, M. and Li, H., 2013. Analysis of complete genomes of *Propionibacterium acnes* reveals a novel plasmid and increased pseudogenes in an acne associated strain. *BioMed research international*, *2013*.
72. Kiatpapan, P., Hashimoto, Y., Nakamura, H., Piao, Y.Z., Ono, H., Yamashita, M. and Murooka, Y., 2000. Characterization of pRGO1, a plasmid from *Propionibacterium acidipropionici*, and its use for development of a host-vector system in propionibacteria. *Applied and environmental microbiology*, *66*(11), pp.4688-4695.
73. Koonin, E.V. and Makarova, K.S., 2013. CRISPR-Cas: evolution of an RNA-based adaptive immunity system in prokaryotes. *RNA biology*, *10*(5), pp.679-686.
74. Koren, S., Harhay, G.P., Smith, T.P., Bono, J.L., Harhay, D.M., Mcvey, S.D., Radune, D., Bergman, N.H. and Phillippy, A.M., 2013. Reducing assembly complexity of microbial genomes with single-molecule sequencing. *Genome biology*, *14*(9), p.R101.
75. Koskinen, P., Deptula, P., Smolander, O.P., Tamene, F., Kammonen, J., Savijoki, K., Paulin, L., Piironen, V., Auvinen, P. and Varmanen, P., 2015. Complete genome sequence of *Propionibacterium freudenreichii* DSM 20271 T. *Standards in genomic sciences*, *10*(1), p.83.
76. Kumar, S.S., Chouhan, R.S. and Thakur, M.S., 2010. Trends in analysis of vitamin B 12. *Analytical biochemistry*, 398(2), pp.139-149.
77. LeBlanc, J.G., Laiño, J.E., del Valle, M.J., Vannini, V., Van Sinderen, D., Taranto, M.P., de Valdez, G., de Giori, G.S. and Sesma, F., 2011. B-Group vitamin production by lactic acid bacteria–current knowledge and potential applications. *Journal of Applied Microbiology, 111*(6), pp.1297-1309.
78. Lee, H., Gurtowski, J., Yoo, S., Nattestad, M., Marcus, S., Goodwin, S., McCombie, W.R. and Schatz, M., 2016. Third-generation sequencing and the future of genomics. *bioRxiv*, p.048603.
79. Letunic, I. and Bork, P., 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic acids research*, *44*(W1), pp.W242-W245.
80. Leuschner, R.G., Robinson, T.P., Hugas, M., Cocconcelli, P.S., Richard-Forget, F., Klein, G., Licht, T.R., Nguyen-The, C., Querol, A., Richardson, M. and Suarez, J.E., 2010. Qualified presumption of safety (QPS): a generic risk assessment approach for biological agents notified to the European Food Safety Authority (EFSA). *Trends in Food Science & Technology*, *21*(9), pp.425-435.
81. Lima-Mendez, G., Van Helden, J., Toussaint, A. and Leplae, R., 2008. Prophinder: a computational tool for prophage prediction in prokaryotic genomes. *Bioinformatics*, *24*(6), pp.863-865.
82. Lingens, B., Schild, T. A., Vogler, B., & Renz, P., 1992. Biosynthesis of vitamin B12: transformation of riboflavin 2H-labeled in the 1'R position or 1'S position into 5, 6-dimethylbenzimidazole. European Journal of Biochemistry, 207, 981e985.
83. Loux, V., Mariadassou, M., Almeida, S., Chiapello, H., Hammani, A., Buratti, J., Gendrault, A., Barbe, V., Aury, J.M., Deutsch, S.M. and Parayre, S., 2015. Mutations and genomic islands can explain the strain dependency of sugar utilization in 21 strains of *Propionibacterium freudenreichii*. *BMC genomics*, *16*(1), p.296.
84. Makarova, K.S., Aravind, L., Wolf, Y.I. and Koonin, E.V., 2011. Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biology direct*, *6*(1), p.38.
85. Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., Barrangou, R., Brouns, S.J., Charpentier, E., Haft, D.H. and Horvath, P., 2015. An updated evolutionary classification of CRISPR-Cas systems. *Nature Reviews Microbiology*. Sep. 28, 2015
86. Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C.J., Lu, S., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R. and Gwadz, M., 2016. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Research*, 45(D), 200-203.
87. Martens, J. H., Barg, H., Warren, M. J., and Jahn, D. (2002). Microbial production of vitamin B12. *Applied Microbiology and Biotechnology*. 58, 275–285.

88. Marwaha, S.S., Sethi, R.P. and Kennedy, J.F., 1983. Influence of 5, 6-dimethylbenzimidazole (DMB) on vitamin B12 biosynthesis by strains of *Propionibacterium*. *Enzyme and Microbial Technology*, *5*(5), pp.361-364.
89. McCoy, R.C., Taylor, R.W., Blauwkamp, T.A., Kelley, J.L., Kertesz, M., Pushkarev, D., Petrov, D.A. and Fiston-Lavier, A.S., 2014. Illumina TruSeq synthetic long-reads empower de novo assembly and resolve complex, highly-repetitive transposable elements. *PloS one*, *9*(9), p.e106689.
90. Meile, L., Le Blay, G. and Thierry, A., 2008. Safety assessment of dairy microorganisms: *Propionibacterium* and *Bifidobacterium*. *International journal of food microbiology, 126*(3), pp.316-320.
91. Menon, I.A. and Shemin, D., 1967. Concurrent decrease of enzymic activities concerned with the synthesis of coenzyme B12 and of propionic acid in propionibacteria. *Archives of biochemistry and biophysics*, *121*(2), pp.304-310.
92. Meurice, G., Jacob, D., Deborde, C., Chaillou, S., Rouault, A., Leverrier, P., Jan, G., Thierry, A., Maillard, M.B., Amet, P. and Lalande, M., 2004. Whole genome sequencing project of a dairy *Propionibacterium freudenreichii* subsp. *shermanii* genome: progress and first bioinformatic analysis. *Le Lait*, *84*(1-2), pp.15-24.
93. Mikheyev, A.S. and Tin, M.M., 2014. A first look at the Oxford Nanopore MinION sequencer. *Molecular ecology resources*, *14*(6), pp.1097-1102.
94. Molina, V., Médici, M., de Valdez, G.F. and Taranto, M.P., 2012. Soybean-based functional food with vitamin B 12-producing lactic acid bacteria. *Journal of Functional Foods*, *4*(4), pp.831-836.
95. Moore, S.J., Lawrence, A.D., Biedendieck, R., Deery, E., Frank, S., Howard, M.J., Rigby, S.E. and Warren, M.J., 2013. Elucidation of the anaerobic pathway for the corrin component of cobalamin (vitamin B12). *Proceedings of the National Academy of Sciences*, *110*(37), pp.14906-14911.
96. Murakami K.A., Hashimoto Y.O., Murooka Y.O. (1993). Cloning and characterization of the gene encoding glutamate 1-semialdehyde 2, 1-aminomutase, which is involved in delta-aminolevulinic acid synthesis in *Propionibacterium freudenreichii. Applied and environmental microbiology. 59*(1):347-50.
97. Murooka, Y., Piao, Y., Kiatpapan, P. and Yamashita, M., 2005. Production of tetrapyrrole compounds and vitamin B12 using genetically engineering of *Propionibacterium freudenreichii*. An overview. *Le Lait*, *85*(1-2), pp.9-22.
98. Nahvi, A., Barrick, J.E. and Breaker, R.R., 2004. Coenzyme B12 riboswitches are widespread genetic control elements in prokaryotes. *Nucleic acids research*, *32*(1), pp.143-150.
99. National Institutes of Health, 2016. Office of Dietary Supplements, Vitamin B12, Dietary Supplement Fact Sheet, Updated: February 11, 2016, accessed 21.3.2017 at: https://ods.od.nih.gov/factsheets/VitaminB12-HealthProfessional/#en5.
100. Nielsen, M.J., Rasmussen, M.R., Andersen, C.B., Nexø, E. and Moestrup, S.K., 2012. Vitamin B12 transport from food to the body's cells—a sophisticated, multistep pathway. Nature Reviews Gastroenterology and Hepatology, 9(6), pp.345-354.
101. O'conell Motherway, M., Zomer, A., Leahy, S.C., Reunanen, J., Bottacini, F., Claesson, M.J., O'Brien, F., Flynn, K., Casey, P.G., Munoz, J.A.M. and Kearney, B., 2011. Functional genome analysis of *Bifidobacterium breve* UCC2003 reveals type IVb tight adherence (Tad) pili as an essential and conserved host-colonization factor. *Proceedings of the National Academy of Sciences*, *108*(27), pp.11217-11222.
102. Obeid, R., Fedosov, S.N. and Nexo, E., 2015. Cobalamin coenzyme forms are not likely to be superior to cyano-and hydroxyl-cobalamin in prevention or treatment of cobalamin deficiency. *Molecular nutrition & food research*, *59*(7), pp.1364-1372.
103. Oh-hama, T., Stolowich, N.J. and Scott, A.I., 1993. 5-Aminolevulinic acid biosynthesis in *Propionibacterium shermanii* and *Halobacterium salinarium*: Distribution of the two pathways of 5-aminolevulinic acid biosynthesis in prokaryotes. *The Journal of General and Applied Microbiology*, *39*(5), pp.513-519.
104. Öhman, T., Lietzén, N., Välimäki, E., Melchjorsen, J., Matikainen, S. and Nyman, T.A., 2010. Cytosolic RNA recognition pathway activates 14-3-3 protein mediated signaling and caspase-dependent disruption of cytokeratin network in human keratinocytes. *Journal of proteome research*, *9*(3), pp.1549-1564.
105. Ojala, T., Laine, P.K., Ahlroos, T., Tanskanen, J., Pitkänen, S., Salusjärvi, T., Kankainen, M., Tynkkynen, S., Paulin, L. and Auvinen, P., 2017. Functional genomics provides insights into the role of *Propionibacterium freudenreichii* ssp. *shermanii* JS in cheese ripening. *International Journal of Food Microbiology*, *241*, pp.39-48.

106. Page, A.J., Cummins, C.A., Hunt, M., Wong, V.K., Reuter, S., Holden, M.T., Fookes, M., Falush, D., Keane, J.A. and Parkhill, J., 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics*, *31*(22), pp.3691-3693.
107. Parizzi, L.P., Grassi, M.C.B., Llerena, L.A., Carazzolle, M.F., Queiroz, V.L., Lunardi, I., Zeidler, A.F., Teixeira, P.J., Mieczkowski, P., Rincones, J. and Pereira, G.A., 2012. The genome sequence of *Propionibacterium acidipropionici* provides insights into its biotechnological and industrial potential. *BMC genomics*, *13*(1), p.562.
108. Patrick, S. and McDowell, A. 2012 Genus I. Propionibacterium in Bergey's Manual of Systematic Bacteriology, Volume Five, The Actinobacteria Part A and B, Second Edition, *Springer Science & Business Media*.
109. Pawlak, R., Parrott, S.J., Raj, S., Cullum-Dugan, D. and Lucus, D., 2013. How prevalent is vitamin B12 deficiency among vegetarians? *Nutrition reviews, 71*(2), pp.110-117.
110. Perez Chaia, A., Sesma, F., Pesce de Ruiz Holgado, A. and Oliver, G., 1988. Screening of plasmids in strains of propionibacterium and mesophilic lactobacilli isolated from Swiss type cheeses. *Microbiologie, aliments, nutrition= Microbiology, foods and feeds, nutrition*.
111. Perignon, M., Vieux, F., Soler, L.G., Masset, G. and Darmon, N., 2017. Improving diet sustainability through evolution of food choices: review of epidemiological studies on the environmental impact of diets. *Nutrition Reviews, 75*(1), pp.2-17.
112. Perlman, D. and Barrett, J.M., 1958. Biosynthesis of cobalamin analogues by *Propionibacterium arabinosum*. *Canadian journal of microbiology*, *4*(1), pp.9-15.
113. Petříček, M., Petříčková, K., Havlíček, L. and Felsberg, J., 2006. Occurrence of two 5-aminolevulinate biosynthetic pathways in *Streptomyces nodosus* subsp. *asukaensi*s is linked with the production of asukamycin. *Journal of bacteriology*, *188*(14), pp.5113-5123.
114. Piao, Y., Kiatpapan, P., Yamashita, M. and Murooka, Y., 2004. Effects of expression of hemA and hemB genes on production of porphyrin in *Propionibacterium freudenreichii*. *Applied and environmental microbiology*, *70*(12), pp.7561-7566.
115. Piveteau, P., 1999. Metabolism of lactate and sugars by dairy propionibacteria: a review. *Le Lait*, *79*(1), pp.23-41.
116. Plyusnin, I., Holm, L. and Kankainen, M., 2009. LOCP—locating pilus operons in gram-positive bacteria. *Bioinformatics*, *25*(9), pp.1187-1188.
117. Polaski, J.T., Holmstrom, E.D., Nesbitt, D.J. and Batey, R.T., 2016. Mechanistic Insights into Cofactor-Dependent Coupling of RNA Folding and mRNA Transcription/Translation by a Cobalamin Riboswitch. Cell reports, 15(5), pp.1100-1110.
118. Quesada-Chanto, A., Schmid-Meyer, A.C., Schroeder, A.G., Fuchter, A., Carvalho-Jonas, M.F., Koehntopp, P.I. and Jonas, R., 1998. Comparison of methods for determination of vitamin B12 in microbial material. *Biotechnology techniques*, *12*(1), pp.75-77.
119. Quick, J., Loman, N.J., Duraffour, S., Simpson, J.T., Severi, E., Cowley, L., Bore, J.A., Koundouno, R., Dudas, G., Mikhail, A. and Ouédraogo, N., 2016. Real-time, portable genome sequencing for Ebola surveillance. *Nature*, *530*(7589), pp.228-232.
120. Rehberger, T.G. and Glatz, B.A., 1990. Characterization of *Propionibacterium* plasmids. *Applied and environmental microbiology*, *56*(4), pp.864-871.
121. Renz P., 1998. Investigations on the Biosynthesis of the 5,6-dimethylbenzimidazole Moiety of Vitamin B12. In: Krautler B., Golding B.T. and Arigoni D. Vitamin B12 and B12 Proteins. WILEY, VCH Verlag GmbH. p 119-130.
122. Renz, P. 1999. Biosynthesis of the 5,6-dimethylbenzimidazole moiety of cobalamin and of the other bases found in natural corrinoids, p. 557–575. In R. Banerjee (ed.), Chemistry and biochemistry of B12. John Wiley & Sons, Inc., New York, NY.
123. Renz, P. and Weyhenmeyer, R., 1972. Biosynthesis of 5, 6-dimethylbenzimidazole from riboflavin Transformation of C-1′ of riboflavin into C-2 of 5, 6-dimethylbenzimidazole. *FEBS letters*, *22*(1), pp.124-126.
124. Reynolds, E., 2006. Vitamin B12, folic acid, and the nervous system. *The lancet neurology*, *5*(11), pp.949-960.
125. Rhoads, A. and Au, K.F., 2015. PacBio sequencing and its applications. *Genomics, proteomics & bioinformatics*, *13*(5), pp.278-289.
126. Richer, M. and Rosselló-Móra, R, 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci,* Nov 10; 106(45):19126-31. doi: 10.1073/pnas.0906412106
127. Rickes, E.L., Brink, N.G., Koniuszy, F.R., Wood, T.R. and Folkers, K., 1948a. Crystalline vitamin B12. *Science (Washington)*, *107*, pp.396-397.

128. Rickes, E.L., Brink, N.G., Koniuszy, F.R., Wood, T.R. and Folkers, K., 1948b. Comparative data on vitamin B12 from liver and from a new source, *Streptomyces griseus*. *Science (Washington)*, *108*, pp.634-635.
129. Roberts, R.J., Vincze, T., Posfai, J. and Macelis, D., 2015. REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. Nucleic acids research, 43(D1), pp.D298-D299.
130. Roessner, C.A., Huang, K.X., Warren, M.J., Raux, E. and Scott, A.I., 2002. Isolation and characterization of 14 additional genes specifying the anaerobic biosynthesis of cobalamin (vitamin B12) in *Propionibacterium freudenreichii* (*P. shermanii*). *Microbiology*, *148*(6), pp.1845-1853.
131. Roth, J.R., Lawrence, J.G., Rubenfield, M., Kieffer-Higgins, S. and Church, G.M., 1993. Characterization of the cobalamin (vitamin B12) biosynthetic genes of *Salmonella typhimurium*. *Journal of bacteriology*, *175*(11), pp.3303-3316.
132. Santos, F., Vera, J.L., Lamosa, P., de Valdez, G.F., de Vos, W.M., Santos, H., Sesma, F. and Hugenholtz, J., 2007. Pseudovitamin B12 is the corrinoid produced by *Lactobacillus reuteri* CRL1098 under anaerobic conditions. *FEBS letters*, *581*(25), pp.4865-4870.
133. Sattler, I., Roessner, C.A., Stolowich, N.J., Hardin, S.H., Harris-Haller, L.W., Yokubaitis, N.T., Murooka, Y., Hashimoto, Y. and Scott, A.I., 1995. Cloning, sequencing, and expression of the uroporphyrinogen III methyltransferase cobA gene of *Propionibacterium freudenreichii (shermanii)*. *Journal of bacteriology*, *177*(6), pp.1564-1569.
134. Savitskaya, E.E., Musharova, O.S. and Severinov, K.V., 2016. Diversity of CRISPR-Cas-mediated mechanisms of adaptive immunity in prokaryotes and their application in biotechnology. *Biochemistry (Moscow)*, *81*(7), pp.653-661.
135. Scholz, C.F. and Kilian, M., 2016. The natural history of cutaneous propionibacteria, and reclassification of selected species within the genus *Propionibacterium* to the proposed novel genera *Acidipropionibacterium gen. nov., Cutibacterium gen. nov*. and *Pseudopropionibacterium gen. nov*. *International Journal of Systematic and Evolutionary Microbiology, 66*(11), pp.4422-4432.
136. Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, 30, pp. 2068-2069.
137. Speedie, D. J. and Hull, G.W., 1960. Cobalamin producing fermentation process. Distillers Co Yeast Ltd, U.S. Patent 2,951,017.
138. Stabler, S.P., 2013. Vitamin B12 deficiency. *New England Journal of Medicine*, *368*(2), pp.149-160.
139. Stierli, M.P., 2002. *DNA transformation of propionibacteria based on plasmids pLME106 and pLME108* (Doctoral dissertation) ETH Zürich, Nr. 14468, 2002.
140. Stupperich, E., and Nexø, E. 1991. Effect of the cobalt-N coordination on the cobamide recognition by the human vitamin B12 binding proteins intrinsic factor, transcobalamin and haptocorrin. *European Journal of Biochemistry*, *199*(2), pp.299-303.
141. Sullivan, M.J., Petty, N.K. and Beatson, S.A., 2011. Easyfig: a genome comparison visualizer. *Bioinformatics*, *27*(7), pp.1009-1010.
142. Taga, M.E. and Walker, G.C., 2008. Pseudo-B12 joins the cofactor family. *Journal of bacteriology*, *190*(4), pp.1157-1159.
143. Taga, M.E., Larsen, N.A., Howard-Jones, A.R., Walsh, C.T. and Walker, G.C., 2007. BluB cannibalizes flavin to form the lower ligand of vitamin B12. *Nature*, *446*(7134), pp.449-453.
144. Taranto, M.P., Vera, J.L., Hugenholtz, J., De Valdez, G.F. and Sesma, F., 2003. *Lactobacillus reuteri* CRL1098 produces cobalamin. *Journal of bacteriology*, *185*(18), pp.5643-5647.
145. Terraf, L., Mendoza, L.M., Tomás, J., Silva, C. and Nader-Macías, M.E.F., 2014. Phenotypic surface properties (aggregation, adhesion and biofilm formation) and presence of related genes in beneficial vaginal lactobacilli. *Journal of applied microbiology*, *117*(6), pp.1761-1772.
146. Thierry, A., Deutsch, S.M., Falentin, H., Dalmasso, M., Cousin, F.J. and Jan, G., 2011. New insights into physiology and metabolism of *Propionibacterium freudenreichii*. *International journal of food microbiology, 149*(1), pp.19-27.
147. Treangen, T.J., Abraham, A.L., Touchon, M. and Rocha, E.P., 2009. Genesis, effects and fates of repeats in prokaryotic genomes. *FEMS microbiology reviews*, *33*(3), pp.539-571.
148. Turroni, F., Serafini, F., Foroni, E., Duranti, S., Motherway, M.O.C., Taverniti, V., Mangifesta, M., Milani, C., Viappiani, A., Roversi, T. and Sánchez, B., 2013. Role of sortase-dependent pili of *Bifidobacterium bifidum* PRL2010 in modulating bacterium–host interactions. *Proceedings of the National Academy of Sciences*, *110*(27), pp.11151-11156.
149. USDA, 2016. US Department of Agriculture, Agricultural Research Service. 2015. Vitamin B12 content USDA National Nutrient Database for Standard Reference Release 28 Internet: http://www.ars.usda.gov/ba/bhnrc/ndl
150. van Niel C.B., 1928. The propionic acid bacteria. Haarlem: Uitgeverszaak and Boissevain and Co.

151. Van Wyk, J., Witthuhn, R.C. and Britz, T.J., 2011. Optimisation of vitamin B 12 and folate production by *Propionibacterium freudenreichii* strains in kefir. *International dairy journal*, *21*(2), pp.69-74.
152. Verderber, E., Lucast, L.J., Van Dehy, J.A., Cozart, P., Etter, J.B. and Best, E.A., 1997. Role of the hemA gene product and delta-aminolevulinic acid in regulation of *Escherichia coli* heme synthesis. *Journal of bacteriology*, *179*(14), pp.4583-4590.
153. Vitreschak, A.G., Rodionov, D.A., Mironov, A.A. and Gelfand, M.S., 2003. Regulation of the vitamin B12 metabolism and transport in bacteria by a conserved RNA structural element. *Rna*, *9*(9), pp.1084-1097.
154. Vorobjeva, L. I. 1999a. Chapter 1: The genus Propionibacterium. Dordrecht, The Netherlands: Kluwer Academic Publishers. pp. 4-46.
155. Vorobjeva, L. I. 1999b. Chapter 7: Economic and medical applications in: Propionibacteria. Dordrecht, The Netherlands: Kluwer Academic Publishers. pp. 209-248
156. Vorobjeva, L. I. 1999c. Chapter 4: Biosynthetic processes and physiologically active compounds. Dordrecht, The Netherlands: Kluwer Academic Publishers. pp. 128-173
157. Warren, M.J., Raux, E., Schubert, H.L. and Escalante-Semerena, J.C., 2002. The biosynthesis of adenosylcobalamin (vitamin B12). *Natural product reports*, *19*(4), pp.390-412.
158. Watanabe, F., Yabuta, Y., Tanioka, Y. and Bito, T., 2013. Biologically active vitamin B12 compounds in foods for preventing deficiency among vegetarians and elderly subjects. Journal of agricultural and food chemistry, 61(28), pp.6769-6775.
159. Wiedenheft, B., van Duijn, E., Bultema, J.B., Waghmare, S.P., Zhou, K., Barendregt, A., Westphal, W., Heck, A.J., Boekema, E.J., Dickman, M.J. and Doudna, J.A., 2011. RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proceedings of the National Academy of Sciences*, *108*(25), pp.10092-10097.
160. Zappa S., Li K., Bauer C.E. 2010. The Tetrapyrrole Biosynthetic Pathway and Its Regulation in *Rhodobacter capsulatus*. In: Hallenbeck P. (eds) Recent Advances in Phototrophic Prokaryotes. Advances in Experimental Medicine and Biology, vol 675. Springer, New York, NY
161. Zhang, Z., Schwartz, S., Wagner, L. and Miller, W. (2000), A greedy algorithm for aligning DNA sequences, J Comput Biol 2000; 7(1-2):203-14